

Tesis Doctoral

APORTE AL PROBLEMA DE LA CORRESPONDENCIA EN ESTEREOSCOPÍA
DE MÚLTIPLES VISTAS EN IMÁGENES AÉREAS

Jean Pierre DÍAZ PAZ

Documento presentado a
Programa de Posgrado en Ingeniería Eléctrica y Electrónica
como requisito para optar por el título de
Doctor en Ingeniería

Director: Humberto LOAIZA CORREA Ph.D.

Escuela de Ingeniería Eléctrica y Electrónica
Facultad de Ingeniería



Agosto 2019

Resumen

La puesta en correspondencia consiste en la identificación de puntos homólogos en distintas imágenes. Esta búsqueda se realiza en 3 etapas básicas: i- detección, en la que se encuentran los puntos característicos de cada imagen; ii- descripción, donde se asigna un vector de características a cada punto detectado; y iii- correspondencia, que se encarga de determinar los puntos homólogos en el espacio de características. En el contexto de la fotogrametría aérea, debido a fenómenos ópticos como variaciones en la dirección de captura y la iluminación, son apropiados los detectores afines y los descriptores locales basados en histogramas de gradientes. Estos detectores y descriptores suelen tomar información de las imágenes en escala de grises y tienen geometrías de muestreo definidas. En esta investigación, se desarrolló un espacio de color no perceptual basado en coordenadas esféricas llamado RTP por las siglas de **R**ho, **T**heta y **P**hi que: desacopla la información de cromaticidad e intensidad, representa las cromaticidades de manera continua y permite aproximar las diferencias cromáticas como distancias euclidianas. Para evaluar el aporte del espacio propuesto sobre la etapa de detección basada en MSER y de la etapa de descripción basada en histogramas de gradientes, se usó un protocolo de pruebas que analiza criterios de *Repetibilidad* y *Razón de verdaderas correspondencias* mediante los valores $F_{0.5}$. Los resultados obtenidos permiten afirmar, con un 95% de certeza, que el uso del espacio de color RTP representa mejoras con diferencias estadísticamente significativas en el desempeño de correspondencia en comparación con otros espacios de color del estado del arte.

Palabras clave— Correspondencia, fotogrametría aérea, color.

Abstract

Image matching allows identifying homologous points on different images. This problem is solved by three basic stages: i- detection, that finds feature points at every image; ii- description, that assigns a features vector to each detected point; and iii- matching, that identifies homologous points at features space. In photogrammetry, affine detectors and local descriptors that are based on histogram of gradients are oftenly used because of different optical conditions such as viewpoint changes, scale changes, compression and illumination. Those detectors and descriptors work on gray scale images and have a fixed geometry. A nonperceptual color space based on spherical coordinates called *RTP* (Rho, Theta, Phi) that decouples chromaticity (θ, φ) and intensity (ρ) information while allowing for a continuous representation of colors and an approximation of chromatic differences as euclidian distances was developed. To assess the effects of RTP on matching, using MSER as detector and descriptors based on histogram of gradients, a test protocol that analyzes Repeteability, Recall and $F_{0.5}$ was used. Results show that using RTP enhances the performance when compared to other state-of-the-art color spaces.

Keywords— Matching, Aerial photogrammetry, Color.

Dedicatoria

Le dedico esta investigación a mis padres Álvaro e Irma.

A mi novia Marcela que me tuvo paciencia.

A mis tíos Nazly, César y Rodrigo por ser un apoyo.

A mi hermano y cuñada Fran y Claudia.

A mi primera sobrina Lauren que ya casi nace.

A mis sobrinas grandes Ana, Johana, Nía, Aleja y Alexa.

A mis compadres Paula, Carlos, Andrea, Danny, Yayis y Fabio por adoptarme en la pasantía.

A mi director Humberto que me tuvo fe desde que era peludo.

A mi profesora Doris que me enseñó matemáticas con sabor.

A VAPS-PSI: La ranga, Blado, Agus, W., George, Ferrín, la paisa, los chuchaquis y los demás que daban bullying, lora, palo y aportes en todos y cada uno de los diferentes aspectos.

A la Universidad, al Valle, a Colombia y a dios.

Agradecimientos

Agradecimientos especiales a los integrantes del Laboratorio de Visión Artificial del grupo de investigación Percepción y Sistemas Inteligentes (PSI) de la Universidad del Valle y a los integrantes del Laboratorio Active Robotic Sensing (ARoS) de North Carolina State University cuyos comentarios y estilos de investigación orientaron el curso de esta investigación.

Este trabajo de investigación se desarrolló con recursos aportados por:

- Departamento Colombiano de Ciencia, Tecnología e Innovación COLCIENCIAS mediante la convocatoria 727 de 2015.
- Universidad del Valle mediante el programa de Asistencia de Docencia 2014-2015 y 2018.

Glosario

ANOVA Análisis de varianza unifactorial. 4, 50, 53, 54, 56, 66, 67

CIELAB CIE 1976 L^* , a^* , b^* . 23

CIELUV CIE L^* , u^* , v^* . 23

GLOH Gradient Location Orientation Histogram. 10

HSI H= Hue, S= Saturation, I= Intensity. 4, 11, 14, 23, 25–27, 29, 30, 33, 35–38, 40, 43–46, 48–53, 55–57, 59, 60, 62, 63, 65–67

HSR Hue Saturation and Ratio (Versión en coordenadas esféricas de HSI). 26, 40, 46, 48–53, 55–57, 67

HSV Hue Saturation and Value. 11, 23, 25–27, 29, 30, 35–38

IoU Intersección sobre Unión. 18, 19

LSIM Least Square Image Matching. 7

MSER Maximally Stable Extremal Regions. 3, 9, 13, 14, 38, 43–46, 48, 49, 51, 53–55, 57, 61, 62, 65, 67

MVS Estereoscopia de Múltiples Vistas. 7

NCC Normalized Cross Correlation. 7

O₁O₂O₃ Colores oponentes. 4, 23, 28, 30, 40, 43, 46, 48–53, 55, 57, 63, 65–68

RGB Red, Green, Blue. 4, 14, 23–25, 27, 28, 30–34, 36, 37, 40, 46, 48–53, 55–59, 63, 65–68

RLAB RLAB. 23

RTP Rho (ρ), Theta (θ), Phi (φ). 3, 4, 30, 31, 33, 35–38, 40, 43–46, 48–60, 62, 63, 65–69

SAD Sumatoria de diferencias absolutas. 21

SIFT Scale Invariant Feature Transform. 3, 4, 9, 10, 13, 14, 16, 17, 38, 46, 48–50, 54, 55, 57, 59, 60, 63, 66–68

SSD Sumatoria de diferencias al cuadrado. 21

SURF Speeded Up Robust Features. 3, 4, 10, 13, 14, 17, 38, 46, 48–51, 53, 59, 60, 63, 66, 67

YC_bC_r Y= Illuminance I= In phase, Q= Quadrature. 23

YIQ Y= Illuminance, I= In phase Q= Quadrature. 23

TABLA DE CONTENIDO

	Pag.
Resumen	I
Abstract	II
Dedicatoria.....	III
Agradecimientos	IV
Glosario	V
1. Introducción.....	1
1.1 Problema de investigación.....	2
1.2 Objetivo general	2
1.3 Objetivos específicos	3
1.4 Aportes	3
1.5 Organización del documento	4
2. Tendencias de los principales procedimientos para la puesta en correspondencia de puntos característicos	6
2.1 Antecedentes	7
2.2 Discusión	13
3. Marco teórico	14
3.1 Detector MSER	14
3.2 Descriptores basados en gradientes	16
3.3 Evaluación del desempeño en el problema de correspondencia	18
3.3.1 Desempeño de los detectores.....	18
3.3.2 Desempeño de los descriptores.....	19
3.4 Correspondencia	21
3.5 Color.....	22
3.5.1 Espacio RGB	23
3.5.2 Espacios HSV, HSI y HSR	25
3.5.3 Colores oponentes	27
3.5.4 Funciones invariantes a color	28
3.6 Discusión	29

4. Propuesta de un espacio de color para los descriptores basados en histogramas de gradientes	30
4.1 Definición del espacio de color RTP	30
4.2 Distancias cromática HS vs. θ φ	33
4.3 Discusión	37
5. Evaluación de los aportes del espacio de color RTP sobre la correspondencia entre imágenes	40
5.1 Data sets	41
5.2 Evaluación del detector MSER	43
5.3 Desempeño de descriptores con color	46
5.3.1 Aporte de RTP a la descripción con SURF	50
5.3.2 Aporte de RTP a la descripción con SIFT	54
5.4 Construcción de un mosaico con imágenes aéreas	58
5.5 Discusión	61
6. Conclusiones	65
7. Perspectivas de trabajos futuros	69
8. Productos	70
Referencias	71
Apendice. A. Curvas de repetibilidad R en la etapa de detección basada en MSER ..	81
Apendice. B. Curvas de Razón de aciertos vs 1-Precisión en la etapa de descripción basada en SURF	85
Apendice. C. Curvas $F_{0.5}$ promedio en la etapa de descripción basada en SURF ..	93
Apendice. D. Curvas de Razón de aciertos vs 1-Precisión en la etapa de descripción basada en SIFT	97
Apendice. E. Curvas $F_{0.5}$ promedio en la etapa de descripción basada en SIFT ...	105

ÍNDICE DE FIGURAS

FIGURA	Pag.
2.1 Imágenes capturadas desde 3 puntos de vista.....	6
2.2 Mapa mental para el problema de puesta en correspondencia entre imágenes.	12
3.1 Detector MSER	15
3.2 Geometría de algunos descriptores del estado del arte.	16
3.3 Filtros wavelet Haar usados en SURF.....	17
3.4 Detección en dos imágenes diferentes.....	18
3.5 Errores de Traslape.	20
3.6 Curvas ideales de R vs $1 - P$	21
3.7 Distribución de colores en el espacio RGB.	24
3.8 Distribución de colores en el espacio HSI	26
4.1 Definición geométrica del espacio RTP.	32
4.2 Gamut de los espacios (a) RTP y (b) HSI/HSV.....	34
4.3 Relación entre de distancias cromáticas en los espacios de color RGB, HSI/HSV y RTP.	37
4.4 (a) Circulo cromático. (b) Descomposiciones del circulo cromático en los espacios RGB, HSV, HSI y RTP.....	39
5.1 Esquema de la solución propuesta.	41
5.2 Ejemplo de imágenes del <i>data set</i>	42
5.3 Desempeño en detección ante cambios de enfoque (Conjunto Bikes).....	44
5.4 Regiones detectadas en la imagen número 4 del conjunto <i>Bikes</i>	45

5.5	Algunas regiones detectadas en la imagen número 4 del conjunto Bikes. ...	46
5.6	Razón de verdaderas correspondencias vs 1-Precisión para cada par del conjunto <i>Wall</i>	47
5.7	Promedio de $F_{0.5}$ con los descriptores SURF con color para cada par del conjunto <i>Wall</i>	49
5.8	Comparación del desempeño de descriptores con color mediante los valores $F_{0.5}$ promedio por conjunto evaluado para el descriptor SURF. ...	49
5.9	Comparación del desempeño de descriptores con color mediante los valores $F_{0.5}$ promedio por conjunto evaluado para el descriptor SIFT.	50
5.10	Intervalos de confianza de valores medios de $F_{0.5}$ para cada transformación con descriptores SURF.	52
5.11	Intervalos de confianza de valores medios de $F_{0.5}$ para todas las transformaciones evaluadas con el descriptor SURF.	55
5.12	Intervalos de confianza de valores medios de $F_{0.5}$ para cada transformación con descriptores SIFT.	57
5.13	Intervalos de confianza de valores medios de $F_{0.5}$ para todas las transformaciones evaluadas con el descriptor SIFT.	59
5.14	Ruta del vuelo para generar el mosaico.	60
5.15	Pares y verdaderas correspondencias con el descriptor SIFT.	61
5.16	Pares y verdaderas correspondencias con el descriptor SURF.	62
5.17	Mosaico construido con la primera linea de vuelo usando SURF con RTP.	64
A.1	Desempeño de detección ante cambios de puntos de vista en una escena estructurada. (Conjunto Graf)	81
A.2	Desempeño de detección ante cambios de puntos de vista en una escena con texturas. (Conjunto Wall)	81
A.3	Desempeño de detección ante cambios en rotación y escala en una escena con textura. (Conjunto Bark)	82
A.4	Desempeño de detección ante cambios de nitidez en escena estructurada. (Conjunto Bikes)	82

A.5	Desempeño de detección ante cambios de nitidez en escena con texturas. (Conjunto Trees).....	83
A.6	Desempeño de detección ante cambios por compresión. (Conjunto UBC) .	83
A.7	Desempeño de detección ante cambios de iluminación. (Conjunto Leuven)	84
A.8	Desempeño de detección con imágenes aéreas. (Conjunto Univalle)	84
B.1	R vs 1-P para el descriptor SURF en el conjunto Graffiti.	85
B.2	R vs 1-P para el descriptor SURF en el conjunto Wall.	86
B.3	R vs 1-P para el descriptor SURF en el conjunto Bark.	87
B.4	R vs 1-P para el descriptor SURF en el conjunto Bikes.....	88
B.5	R vs 1-P para el descriptor SURF en el conjunto Trees.....	89
B.6	R vs 1-P para el descriptor SURF en el conjunto UBC.	90
B.7	R vs 1-P para el descriptor SURF en el conjunto Leuven.....	91
B.8	R vs 1-P para el descriptor SURF en el conjunto Univalle.....	92
C.1	F para cada par del conjunto Graffiti con el descriptor SURF con color ..	93
C.2	F para cada par del conjunto Wall con el descriptor SURF con color	93
C.3	F para cada par del conjunto Bark con el descriptor SURF con color	94
C.4	F para cada par del conjunto Bikes con el descriptor SURF con color	94
C.5	F para cada par del conjunto Trees con el descriptor SURF con color	95
C.6	F para cada par del conjunto UBC con el descriptor SURF con color	95
C.7	F para cada par del conjunto Leuven con el descriptor SURF con color ..	96
C.8	F para cada par del conjunto Univalle con el descriptor SURF con color .	96
D.1	R vs 1-P para el descriptor SIFT en el conjunto Graffiti.....	97
D.2	R vs 1-P para el descriptor SIFT en el conjunto Wall.....	98
D.3	R vs 1-P para el descriptor SIFT en el conjunto Bark.	99
D.4	R vs 1-P para el descriptor SIFT en el conjunto Bikes.....	100

D.5	R vs 1-P para el descriptor SIFT en el conjunto Trees.....	101
D.6	R vs 1-P para el descriptor SIFT en el conjunto UBC.	102
D.7	R vs 1-P para el descriptor SIFT en el conjunto Leuven.....	103
D.8	R vs 1-P para el descriptor SIFT en el conjunto Univalle.....	104
E.1	F para cada par del conjunto Graffiti con el descriptor SIFT con color ...	105
E.2	F para cada par del conjunto Wall con el descriptor SIFT con color	105
E.3	F para cada par del conjunto Bark con el descriptor SIFT con color	106
E.4	F para cada par del conjunto Bikes con el descriptor SIFT con color	106
E.5	F para cada par del conjunto Trees con el descriptor SIFT con color	107
E.6	F para cada par del conjunto UBC con el descriptor SIFT con color	107
E.7	F para cada par del conjunto Leuven con el descriptor SIFT con color ...	108
E.8	F para cada par del conjunto Univalle con el descriptor SIFT con color ..	108

ÍNDICE DE CUADROS

TABLA		Pag.
4.1	7 Colores representados en tres espacios de color.	34
4.2	Ángulos en el espacio RGB como distancias cromáticas [rad].	35
4.3	Distancias cromáticas calculadas en los espacios HSI/HSV (H, S).	35
4.4	Distancias cromáticas calculadas en el espacio RTP (θ, φ).	36
5.1	Resumen de los valores de repetibilidad con umbral fijo del 40 %.	43
5.2	Resultados estadísticos del análisis ANOVA para el descriptor SURF.	51
5.3	Resultados estadísticos del análisis ANOVA general para el descriptor SURF.	54
5.4	Resultados estadísticos del análisis ANOVA para el descriptor SIFT.	56
5.5	Resultados estadísticos del análisis ANOVA general para el descriptor SIFT.	58
5.6	Resumen de correspondencias encontradas entre dos imágenes aéreas con SIFT.	60
5.7	Resumen de correspondencias encontradas entre dos imágenes aéreas con SURF.	63

1. Introducción

La fotogrametría aérea es una técnica a través de la cual se estiman dimensiones y posiciones de los objetos en el espacio y se obtienen datos para la elaboración de cartografía topográfica y temática en aplicaciones como agronomía, cartografía, ortofotografía, minería, arquitectura, planeamiento del territorio, medio ambiente, arqueología, control de estructuras, topografía, entre otras [Haala, 2011]. Los avances tecnológicos han provocado el desarrollo de cámaras con cada vez mejores prestaciones, a la vez que procesadores con mayor poder computacional que posibilitan el procesamiento de un creciente volumen de datos y el incremento del traslape entre imágenes del 60 %, entre imágenes consecutivas en una línea de vuelo, y 20 %, entre líneas adyacentes, al 80 % o más para aumentar la redundancia favoreciendo la eliminación de datos anómalos, la disminución de los problemas de oclusión y el incremento de la precisión de los resultados [Haala, 2011, Hartley and Zisserman, 2004].

La puesta en correspondencia es la base de los procesos fotogramétricos [Haala, 2011] y una tarea fundamental en el campo de la visión artificial [Szeliski, 2011, Yang et al., 2018] que consiste en la identificación de los puntos que corresponden al mismo elemento en una escena 3D en distintas imágenes. La búsqueda de puntos homólogos requiere de una medida de similitud y una metodología de selección [Hong, 2010]. Esta búsqueda se ha resuelto mediante métodos i- **locales**, que se basan en la semejanza presente entre vecindades medida con métricas que evalúan la similitud entre ventanas determinadas; ii- **globales**, que utilizan información completa de las imágenes para optimizar iterativamente una función de costo generada al asociar disparidades a cada píxel [Brown et al., 2005, Barandiaran et al., 2013]; y iii- **semiglobales**, en los que se inicia con la correspondencia de puntos especiales detectados en las imágenes [Bethmann and Luhmann, 2014] para encontrar una superficie de disparidad inicial que posteriormente se optimiza [Hirschmüller, 2008].

Los métodos semiglobales han mostrado una mejor relación entre el tiempo de procesamiento y la precisión obtenida [Xiao et al., 2013, Hirschmuller, 2005]. Estos métodos constan de 3 etapas básicas [Bay et al., 2008]: i- detección, en la que se

encuentran los puntos característicos de cada imagen; ii- descripción, donde se asigna un vector de características a cada punto detectado; y iii- correspondencia, que se encarga de determinar los puntos homólogos en el espacio de características.

Los errores en correspondencia se presentan debido a que las técnicas utilizadas en las etapas descritas no son totalmente robustas ante variaciones en la dirección de captura, las condiciones de iluminación, distorsiones debidas a la construcción física de las cámaras, vibraciones introducidas desde la plataforma de transporte, etc [Tola et al., 2011, Zhou et al., 2014]. Además, distintos fenómenos ópticos como la oclusión y distorsión por perspectiva dificultan la búsqueda de puntos homólogos en imágenes de entornos no estructurados y no controlados [Tola et al., 2010, Xiao et al., 2013, Correal et al., 2014, Tsay and Lee, 2012], haciendo de la correspondencia un problema de investigación abierto [Bradley et al., 2008, Calderon et al., 2013, Ackermann et al., 2014, Trzcinski et al., 2015, Colomina and Molina, 2014, Zhou and Payandeh, 2015].

1.1 Problema de investigación

Con base en el contexto descrito, se evidencia la necesidad de obtener técnicas de detección y descripción de puntos característicos en imágenes aéreas que sean robustas [Ramírez Rivera et al., 2014] ante variaciones de dirección de captura, condiciones de iluminación, distorsiones propias de las cámaras, y otros fenómenos ópticos, a la vez que ofrezcan facilidades para encontrar puntos homólogos [Calonder et al., 2010], por lo cual la pregunta de investigación abordada en este proyecto fue:

- ¿Cómo se pueden detectar y describir puntos característicos en las imágenes para mejorar el desempeño en la correspondencia basada en descriptores?.

1.2 Objetivo general

- Desarrollar un procedimiento computacional para la detección y descripción de puntos característicos para la puesta en correspondencia de imágenes aéreas en estereoscopia de múltiples vistas.

1.3 Objetivos específicos

- Reconocer las características de los principales procedimientos utilizados para la puesta en correspondencia de puntos característicos en estereoscopía de múltiples vistas.
- Proponer una metodología para la detección y descripción de puntos característicos en múltiples imágenes.
- Evaluar el desempeño de la metodología de puesta en correspondencia propuesta con bases de datos de imágenes aéreas.

1.4 Aportes

En esta investigación se desarrolló un procedimiento para la puesta en correspondencia entre imágenes con métodos semiglobales cuyas etapas de detección y descripción emplean metodologías destacadas de la literatura consultada. En la etapa de detección se emplea el detector MSER y en la etapa de descripción se emplean los descriptores basados en histogramas de gradientes más destacados entre los documentos consultados; SIFT y SURF. Las metodologías de detección y descripción se enriquecieron usando un espacio de color propuesto en esta investigación. El espacio propuesto, llamado RTP por las siglas de **R**ho, **T**heta y **P**hi, es un espacio de color no perceptual, basado en coordenadas esféricas, que desacopla la información de cromaticidad de la relacionada con la intensidad, representa los tonos rojizos de manera continua y permite aproximar las diferencias cromáticas como distancias euclidianas, lo cual mejora la estimación de gradientes y la robustez de los histogramas de gradientes como descriptores.

Para evaluar el aporte del espacio propuesto sobre las etapas de detección y descripción se usó un protocolo de pruebas que analiza criterios de repetibilidad y verdaderas correspondencias [Mikolajczyk et al., 2005]. En este protocolo se realizan los análisis comparativos al evaluar, mediante inspección visual, las curvas de *Razón de verdaderas correspondencias* vs. *1-Precisión* que se obtienen para algunos pares estereoscópicos con transformaciones conocidas [Mikolajczyk et al., 2005]. Para evitar

subjetividades en el análisis de las curvas se usa el valor máximo de la función $F_{0.5}$ calculada sobre las curvas mencionadas como un indicador cuantitativo del desempeño en correspondencia [Lee and Park, 2017]. Para evitar posibles errores derivados de la selección del valor máximo de un indicador como parámetro de comparación, en este trabajo se incorpora un Análisis de varianza unifactorial (ANOVA) sobre los valores $F_{0.5}$ obtenidos con todos los pares estereoscópicos analizados. De esta manera se evalúan los datos completos de las curvas y se determina, con criterios estadísticos, si existen o no diferencias significativas en el desempeño de distintas técnicas de correspondencia.

Los resultados obtenidos permiten afirmar con un 95% de certeza que el uso del espacio de color RTP introduce mejoras con diferencias estadísticamente significativas en el desempeño de correspondencia con el descriptor SURF y mejoras leves con el descriptor SIFT en comparación con los espacios de color RGB, HSI, de colores oponentes $O_1O_2O_3$, otro espacio con coordenadas esféricas, la representación con una función invariante cromática y tonos de gris.

Finalmente, aunque en esta investigación se tomaron los descriptores SIFT y SURF, los resultados obtenidos sugieren que el espacio RTP puede mejorar el desempeño de otros descriptores.

1.5 Organización del documento

Este documento está constituido por 7 capítulos. En el capítulo 1 se presenta el problema de investigación y se introducen los temas más relevantes en el contexto de la correspondencia entre imágenes: detección y descripción. El capítulo 2 muestra los aspectos más relevantes que fueron identificados a partir de una revisión sistemática de literatura científica relacionada con las temáticas. El capítulo 3 provee un marco general de las técnicas, identificadas como las principales en el estado del arte y que soportan la propuesta de esta tesis: detectores de puntos característicos, descriptores basados en histogramas de gradiente y espacios de color dependientes del dispositivo. En el capítulo 4 se describe un espacio de color propuesto para mejorar el desempeño de los descriptores basados en histogramas de gradientes y se realiza un análisis comparativo de las diferencias cromáticas en distintos espacios de color. El capítulo 5

describe la metodología general para evaluar el desempeño en los problemas de puesta en correspondencia y los *data sets* que se utilizaron para la evaluación del desempeño de la metodología, así como los resultados obtenidos en cada una de las pruebas y un análisis cuantitativo de los desempeños obtenidos con distintos espacios de color en el problema de puesta en correspondencia. El capítulo 6 lista las conclusiones y el capítulo 7 lista los trabajos futuros que dan continuidad a este proyecto de investigación.

2. Tendencias de los principales procedimientos para la puesta en correspondencia de puntos característicos

Se hizo una revisión sistemática de documentos relacionados con detectores de puntos característicos y descriptores locales para la puesta en correspondencia entre imágenes. Se seleccionaron los documentos publicados entre enero de 2001 y diciembre de 2018 en los idiomas inglés y español. Esta revisión se hizo usando las bases de datos bibliográficas **SCOPUS**, **ACM-DL** e **IEEEEXPLORE**. A continuación se describen las tendencias identificadas en los campos explorados y se realiza una discusión respecto a las posibles mejoras en el área.

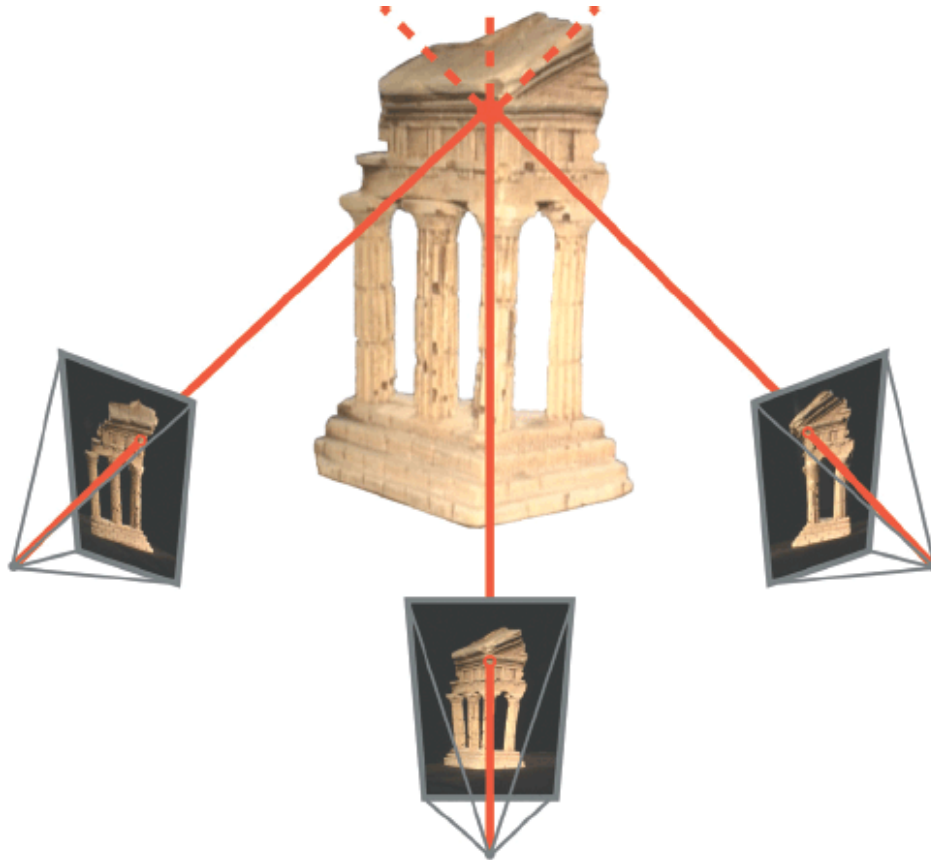


Figura 2.1: Imágenes capturadas desde 3 puntos de vista.
Tomada de [Goesele et al., 2010].

2.1 Antecedentes

La puesta en correspondencia, base de la fotogrametría y la Estereoscopía de Múltiples Vistas (MVS), consiste en la identificación de los puntos generados por la proyección del mismo objeto en una escena 3D sobre un conjunto de imágenes capturadas desde distintos puntos de vista [Szeliski, 2011, Haala, 2011, Lee and Park, 2017], como se ilustra en la Figura 2.1. Los errores en correspondencia se presentan debido a que las metodologías no son totalmente robustas ante variaciones en la dirección de captura, las condiciones de iluminación, distorsiones debidas a la construcción física de las cámaras, vibraciones introducidas desde la plataforma de transporte, etc [Tola et al., 2011, Zhou et al., 2014, Mikolajczyk and Schmid, 2005, Trzcinski et al., 2015]. Además, distintos fenómenos ópticos como la oclusión y distorsión por perspectiva dificultan la búsqueda de puntos homólogos en imágenes de entornos no estructurados y no controlados [Bradley et al., 2008, Tola et al., 2010, Tsay and Lee, 2012, Calderon et al., 2013, Xiao et al., 2013, Correal et al., 2014, Ackermann et al., 2014, Colomina and Molina, 2014, Zhou and Payandeh, 2015].

Los métodos de correspondencia pueden agruparse en 3 familias:

- Los métodos **locales** se basan en la semejanza entre vecindades. La semejanza se estima acumulando una función de costo en la cobertura de una ventana móvil. Las funciones Normalized Cross Correlation (NCC) y Least Square Image Matching (LSIM) son típicas en la estimación de dicha similitud. Esta familia de métodos se fundamenta en el supuesto de disparidad constante dentro de las vecindades de correlación [Hirschmuller, 2005]. Por esta razón debe definirse una vecindad de tamaño suficientemente grande para contener las variaciones necesarias y encontrar correspondencias, y tan pequeña para minimizar las variaciones de disparidad dentro de dichas ventanas [Hong, 2010]. Generalmente se usan ventanas cuadradas, sin embargo se han propuesto técnicas que seleccionan ventanas de verificación [Szeliski, 2001], adaptan dinámicamente las formas y tamaños de la ventana utilizada [Kanade and Okutomi, 1991] o asignan pesos ponderados en las ventanas [Yoon and Kweon, 2006] de acuerdo con propiedades de los píxeles que rodean la zona que está siendo apareada

[Hosni et al., 2013]. La búsqueda de correspondencia por métodos locales ha sido utilizada en aplicaciones en tiempo real gracias a la rapidez de implementación que se obtiene al limitar la zona de búsqueda por la geometría epipolar propia de un sistema de cámaras fijas [Hirschmüller, 2008, Hsia et al., 2012].

- Los métodos **globales** utilizan la información completa de las imágenes para optimizar una función del costo al asociar una disparidad d_p a cada píxel p [Brown et al., 2005, Barandiaran et al., 2013]. Generalmente la función de costo utilizada relaciona los datos y la suavidad de la superficie generada. Estos algoritmos pueden proporcionar resultados más precisos que los métodos locales, pero requieren grandes recursos y complejidad computacionales [Hirschmuller, 2005, Tsay and Lee, 2012]. En el contexto de la estereoscopia de múltiples vistas este enfoque ha sido considerado por distintos autores [Elibol et al., 2008, Li et al., 2015] y se ha usado para optimizar una superficie a partir de un resultado inicial mediante superficies probabilísticas [Calakli et al., 2012].
- En los métodos **semiglobales** se inicia con la correspondencia de puntos especiales identificados en las imágenes [Bethmann and Luhmann, 2014] para encontrar una superficie de disparidad que posteriormente se optimiza [Hirschmüller, 2008]. Se obtiene una superficie inicial a partir de puntos característicos en las imágenes para, a partir de esta superficie, comenzar el proceso iterativo de los métodos globales. Estos métodos han demostrado tener mejor relación entre el tiempo de ejecución y la precisión obtenida [Xiao et al., 2013, Hirschmuller, 2005] que los métodos globales.

Los métodos de correspondencia semiglobales constan de 3 etapas básicas [Bay et al., 2008]: una etapa de **detección** de puntos característicos en las imágenes; una etapa de **descripción** local, en la que se asigna un vector de características a cada punto detectado con base en la información de una vecindad al rededor del punto; una etapa de **correspondencia**, en la que se encuentran grupos de descriptores en el espacio de características.

Los detectores y descriptores deben tener cierta invariancia ante distintas transformaciones geométricas [Mikolajczyk and Schmid, 2003, Lobaton et al., 2011, Szeliski, 2011, Yang et al., 2018] que en el contexto de la fotogrametría aérea,

donde se capturan imágenes de escenas rígidas desde distintos puntos de vista y bajo diferentes condiciones, pueden modelarse como rotaciones, escalamientos anisotrópicos y traslaciones, es decir, transformaciones afines [Li and Zhang, 2012].

De acuerdo con la evaluación del desempeño de detectores y descriptores locales [Mikolajczyk and Schmid, 2003] que posteriormente se convirtió en el estándar para este tipo de evaluaciones [Mikolajczyk and Schmid, 2005, Burghouts and Geusebroek, 2009, Liao and Liu, 2010, Canclini et al., 2013, Lee and Park, 2017], los detectores Harris–Laplace, SIFT [Lowe, 1999] tienen los mejores desempeños ante alteraciones fotométricas pero reducen su desempeño ante los cambios geométricos propios de las capturas desde distintos puntos de vista. Por su parte, los detectores de regiones afines Harris–Affine [Mikolajczyk and Schmid, 2004], Hessian–Affine [Mikolajczyk and Schmid, 2004] y MSER (Maximally Stable Extremal Regions) [Matas et al., 2004] son más apropiados para los problemas de correspondencia en el contexto de la fotogrametría aérea, pues las regiones detectadas, al tener en cuenta las transformaciones mencionadas, se adaptan a las distorsiones geométricas locales de la imagen [Lee and Park, 2017]. Entre los detectores afines, MSER se destaca por tener el mejor desempeño en términos de repetibilidad [Gauglitz et al., 2011].

Los descriptores locales de mayor uso se basan en histogramas de gradientes orientados. Estos encierran la información de cómo se distribuyen las diferencias entre los píxeles con respecto a la orientación y posición relativa en la vecindad del punto descrito [Lowe, 2004, Verma and Liu, 2017a, Yang et al., 2018] y reducen la dependencia ante los cambios en la iluminación de la escena [Szeliski, 2011] al utilizar las diferencias de tonos dentro de las vecindades en lugar de las intensidades. El descriptor local de mayor popularidad es SIFT [Lowe, 2004, Verma and Liu, 2017b]; una concatenación de histogramas de gradientes en geometría cartesiana. Desde que fue propuesto, SIFT ha sido paralelizado [Moreno et al., 2009], se ha implementado sin el uso de etapas de suavizado [Liu et al., 2010], se ha adaptado su geometría [Khan et al., 2011, Liao and Liu, 2010] e incorporado información de contraste en la periferia [Su et al., 2013], se han usado abanicos en las vecindades para lidiar con problemas de discontinuidad en las escenas [Cui and Ngan, 2011], se ha alterado su geometría [Mikolajczyk and Schmid, 2005] mediante transformaciones de perspectiva [Clark et al., 2008] y de coordenadas esféricas [Cruz-Mota et al., 2012] y se ha combinado SIFT con la función de valores de píxel para mejorar su desempeño en la etapa de

correspondencia [Fan et al., 2008], entre otros trabajos que lo han tomado como base para generar nuevos descriptores [Tola et al., 2008, Ke et al., 2004, Mikolajczyk and Schmid, 2005]. Entre los descriptores basados en SIFT que son referente en la literatura científica se destacan SURF [Liu et al., 2011], que aproxima los histogramas a partir de filtrados y sumatorias en geometría cartesiana; GLOH (Gradient Location Orientation Histogram) [Mikolajczyk and Schmid, 2005], que genera histogramas en una geometría log – polar; y DAISY [Tola et al., 2008], que se construye al concatenar muestras de la imagen que se ha procesado con una serie de filtros definidos de acuerdo con una geometría polar.

En los descriptores mencionados anteriormente no se incluye la información de color [Verma and Liu, 2017b], una característica de bajo nivel que brinda información relevante en diversas aplicaciones de visión por computador [Shamoi et al., 2014]. Sin embargo varios autores han mostrado que la información cromática resulta menos sensible al ruido y variaciones fotométricas [Alegre and Pajares, 2016]; lo que permite intuir que la implementación de descriptores con información cromática puede reducir la cantidad de falsas correspondencias [Cai et al., 2005, Huang et al., 2014, van de Sande et al., 2010, Verma and Liu, 2017b]. En estudios anteriores se ha realizado corrección gamma y de color antes de pasar las imágenes a tonos de gris y encontrar los descriptores [Abraham et al., 2013]; se han desarrollado descriptores locales utilizando distintos espacios de color [Verma and Liu, 2017b] al concatenar los descriptores obtenidos en cada componente de un espacio dado [Rassem and Khoo, 2011]; también se ha complementado la información del gradiente de cada píxel incluyendo las magnitudes del gradiente en cada dirección y la orientación, ponderados por gaussianas junto a las componentes en el espacio de color *HSV* para encontrar la matriz de covarianza de las vecindades de los puntos de interés [Lang et al., 2010] como descriptor; se han usado coordenadas cromáticas para estimar valores invariantes de color como una combinación lineal de estas componentes para usar SIFT con los valores de colores invariantes en la formación de un descriptor con color [Abdel-Hakim and Farag, 2006]; además se han usado combinaciones de las representaciones invariantes *rgb*, $l_1l_2l_3$, y matiz para aportar información cromática al descriptor SIFT [Zhang et al., 2007]; se ha usado el espacio de colores oponentes [Liu and Li, 2013] y se ha usado información cromática, de textura y de geometría para describir, usando SIFT las regiones definidas con diagramas de Voronoi [Do et al., 2013].

Un color puede representarse mediante su distribución espectral de potencia en un amplio rango de longitudes de onda, sin embargo los conos de la retina humana captan la luz en solo tres bandas espectrales (Rojo: $700nm$, Verde: $550nm$ y Azul: $440nm$) [Ford and Roberts, 1998]. Debido a esto, se han propuesto diversos espacios de color [Alarcon and Dalmau, 2014], usualmente de tres dimensiones [Susstrunk et al., 1999], para representar la reacción humana ante los estímulos en las tres longitudes de onda mencionadas [Zhou and Hansen, 2016]. Así, la información contenida en las imágenes de color se relaciona directamente con la escena y se codifica de acuerdo con el espacio de color utilizado. Se ha planteado un modelo en el que las componentes de matiz (H) y saturación (S) de los modelos HSI y HSV se expresan como ángulos en un sistema de coordenadas esféricas con el eje polar en la línea de grises dándole a este modelo la intuitividad que ya tienen los modelos HSI y HSV [Chen et al., 2013a, Levene, 1993]. Estudios previos de la percepción del color en primates han mostrado que los colores se distribuyen sobre una esfera $4D$ [Latanov et al., 1991] de radio unitario [Leonov et al., 2008], en una semiesfera $3D$ [Sokolov et al., 1979] o en un octante de esfera [Gómez, 2016].

La Figura 2.2 muestra un mapa mental que hace énfasis en los hallazgos de la revisión realizada. En este mapa se han resaltado con los elementos que, según la literatura, resultan más apropiados y se tomaron como base para la investigación.

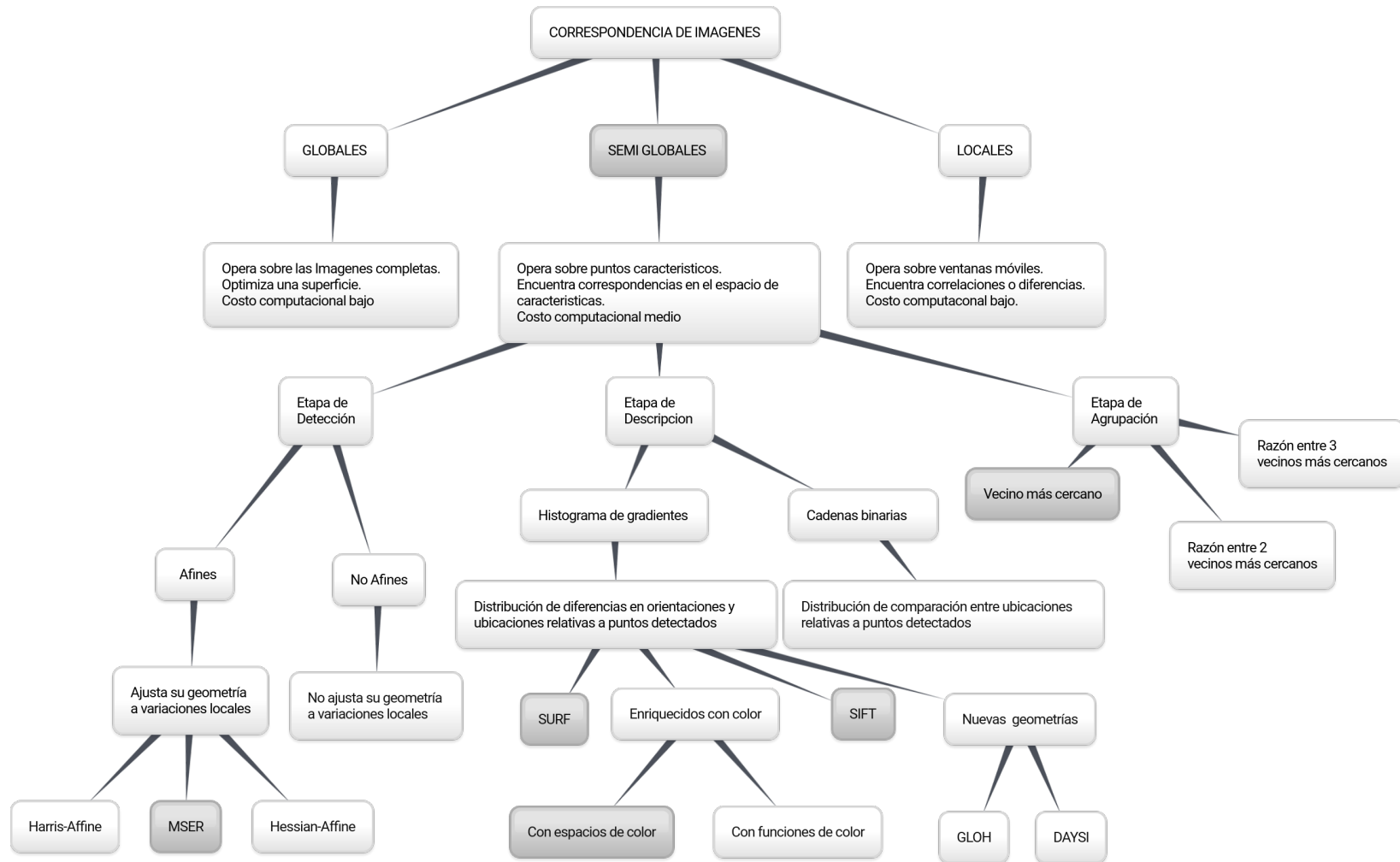


Figura 2.2: Mapa mental para el problema de puesta en correspondencia entre imágenes.

2.2 Discusión

- Se requiere de detectores invariantes ante transformaciones afines para que la información de las vecindades relacionadas con los puntos característicos pueda adaptarse a las transformaciones geométricas ocasionadas por las condiciones de captura propias de un vuelo fotogramétrico. De acuerdo con la literatura consultada, el detector MSER reporta el mejor desempeño en términos de repetibilidad.
- La invariancia de los descriptores locales se soporta en la capacidad de adaptar su geometría a las vecindades de los puntos característicos de manera coherente con las transformaciones debidas a la geometría de captura y en la capacidad de adaptar sus elementos a las transformaciones radiométricas relacionadas con los cambios de iluminación. Los descriptores locales basados en histogramas de gradientes condensan la información de cómo se distribuyen las diferencias de la componente descrita dentro de una vecindad respecto a la orientación. Los descriptores SIFT y SURF se reportan en la literatura consultada como los de mayor uso con los más altos desempeños en términos de *Razón de Verdaderas Correspondencias* y *Precisión*.
- Para el caso de interés de esta investigación, el desempeño de los descriptores puede mejorarse buscando una geometría que se adapte a los requerimientos de la correspondencia entre imágenes aéreas o representando el color de manera que aumente la invariancia ante los distintos fenómenos propios de la fotogrametría.

3. Marco teórico

La propuesta desarrollada en esta investigación se basa en el detector de regiones características MSER y los descriptores basados en histogramas de gradientes operando sobre las componentes de un espacio de color propuesto. En este capítulo se describe el funcionamiento del detector MSER, la descripción de puntos característicos usando descriptores basados en histogramas de gradientes, en particular usando los descriptores SIFT y SURF, y los espacios de color RGB y HSI.

3.1 Detector MSER

MSER detecta regiones intrínsecamente conectadas, que son más oscuras o más brillantes que las regiones que las rodean, y son estables ante transformaciones geométricas y fotométricas. Este detector ha sido identificado como uno de los de mejor desempeño en términos de precisión y repetibilidad [Mikolajczyk and Schmid, 2005]. Las regiones detectadas ofreciendo ciertas propiedades, entre las que se destacan:

- Invarianza ante cambios de intensidad.
- Conservación de la adyacencia entre regiones (en ausencia de oclusión).
- Detección en diversas escalas.

El detector MSER toma una imagen monocromática I para aplicar una secuencia de umbrales. Para cada umbral t_h se encuentran las regiones R tales que $I(R) \geq t_h$ y se halla el área respectiva $A_R(t_h)$. Finalmente, se seleccionan los umbrales que corresponden a mínimos locales de la razón de cambio del área de cada región y se encuentran las regiones MSER representándolas por el punto de menor nivel y el umbral correspondiente [Matas et al., 2004]. La Figura 3.1 muestra el proceso mediante el cual se define una región conectada a un punto de interés en la letra V del logo de la Universidad del Valle. La Figura 3.1a muestra la imagen monocromática

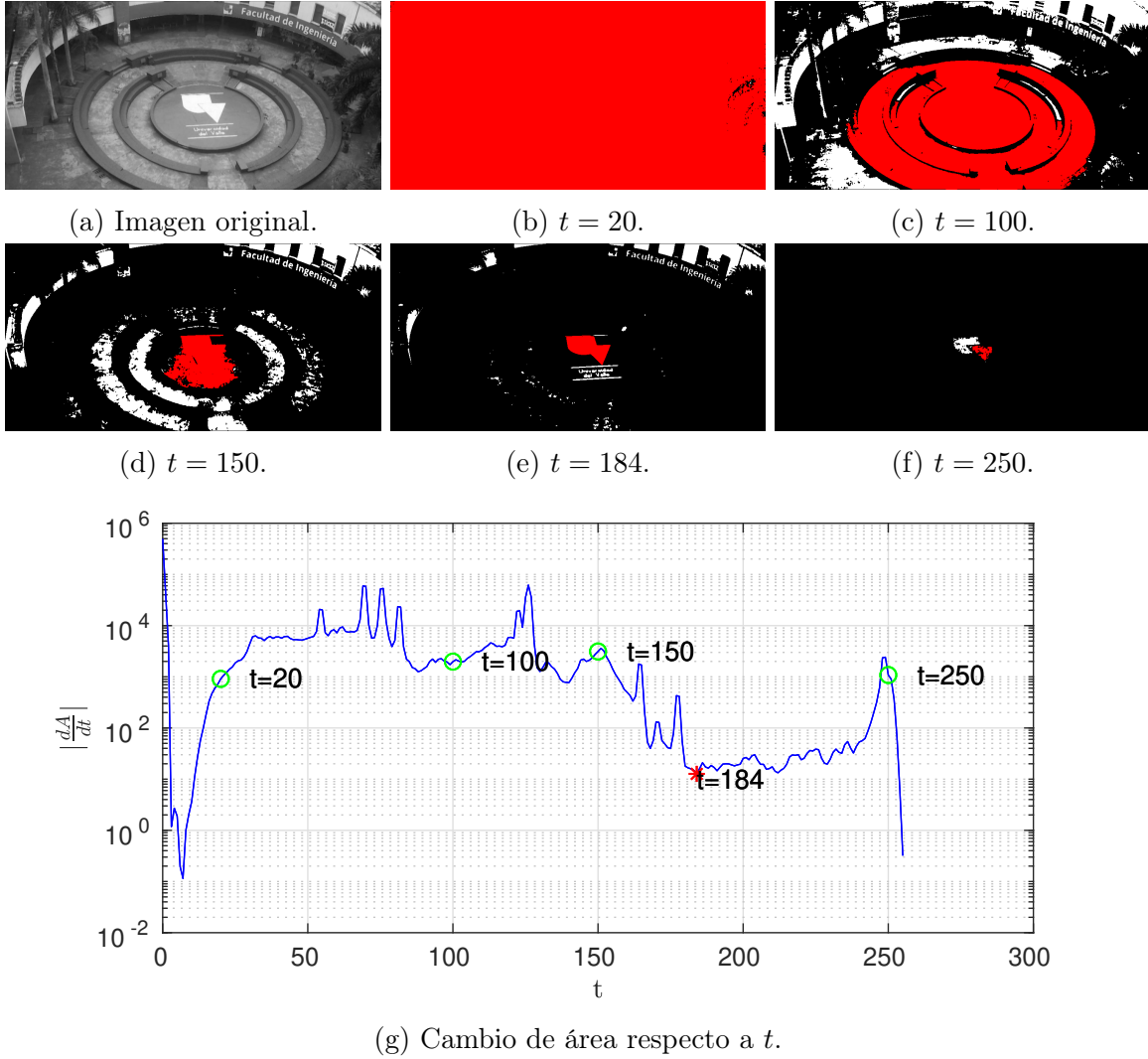


Figura 3.1: Detector MSER

original sobre la que se aplica el detector. En las Figuras 3.1b, 3.1c, 3.1d, 3.1e y 3.1f se resalta en blanco las regiones que cumplen con la umbralización ante los valores $t = \{20, 100, 150, 184 \text{ y } 250\}$ respectivamente, y en rojo la región que además está conectada al punto de interés. La Figura 3.1g ilustra la razón de cambio del área de la región conectada al punto de interés respecto al umbral evaluado. En esta curva se puede observar que $t_h = 184$ es el umbral apropiado para la detección de la región de interés evaluada. Finalmente, se encuentran las elipses que contienen las regiones detectadas y se asignan los centroides con los parámetros de las elipses a los puntos característicos y sus respectivas deformaciones afines.

3.2 Descriptores basados en gradientes

En el contexto de la correspondencia, los gradientes han contribuido a la robustez en la descripción de puntos característicos; pues al utilizar las diferencias en las vecindades en lugar de los valores de gris se disminuye la dependencia ante la iluminación de la escena. Los histogramas de gradientes permiten describir cómo se distribuyen dichas diferencias respecto a la orientación y posición en la vecindad del punto de interés.

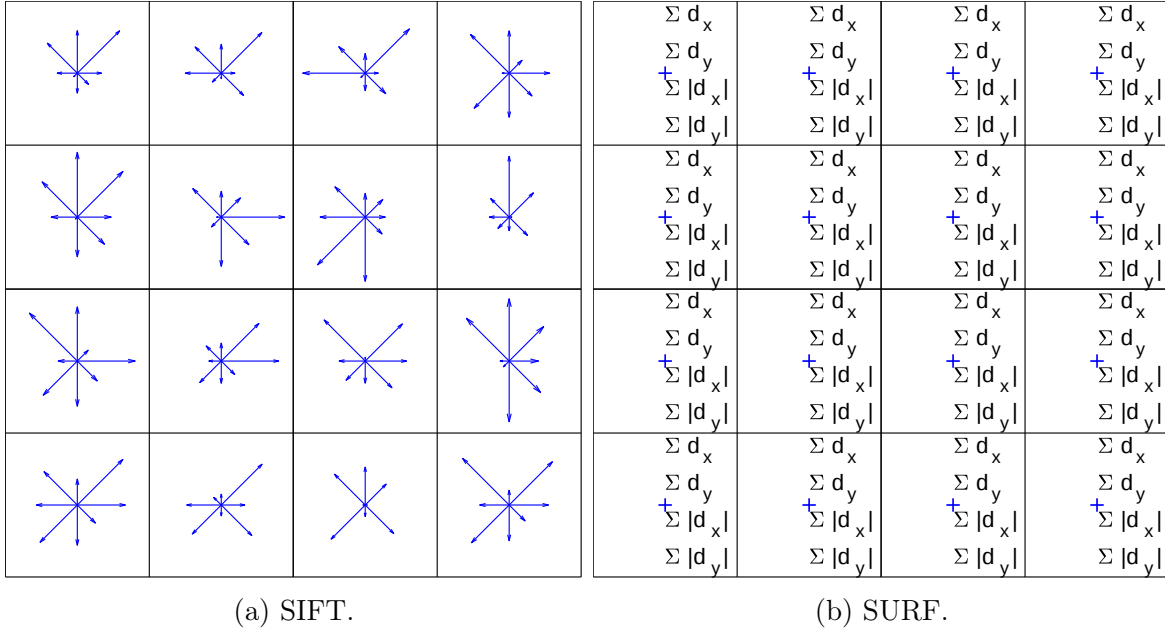


Figura 3.2: Geometría de algunos descriptores del estado del arte.

SIFT [Lowe, 2004] utiliza una aproximación a la primera derivada en dirección horizontal d_x y en dirección vertical d_y . Inicialmente, dada una geometría normalizada que se determina en la etapa de detección, se toma una región cuadrada alrededor del punto de interés detectado. La región es de 16×16 píxeles y orientada según la dirección predominante para el punto característico. Se encuentran los gradientes en los píxeles de la región mediante la convolución de la imagen con los kernels lineales $k_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ y $k_y = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T$. Las magnitudes de los gradientes se ponderan con una Gaussiana para reducir la influencia de los píxeles más alejados al punto característico. La región mencionada se divide en 4×4 sub-regiones y se encuentra, en cada sub-región, un histograma de gradientes con 8 rangos de direcciones de interés ($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$) mediante la suma de las magnitudes de

los gradientes dentro cada rango. Finalmente, el descriptor SIFT se forma concatenando los histogramas encontrados en las 16 sub-regiones. La Figura 3.2a muestra la cuadrícula de 4×4 sub-regiones asociadas a los histogramas de gradientes que generan el descriptor SIFT. El proceso descrito muestra que la diferencia entre píxeles es la base para formar el descriptor.

SURF [Bay et al., 2008] toma una región cuadrada alrededor del punto de interés detectado y rotada según la orientación predominante de dicha región. Inicialmente, la región se divide en 4×4 sub-regiones cuadradas y se computan las respuestas d_x y d_y Haar wavelet (ver Figura 3.3) de cada sub-región en las direcciones horizontal y vertical respectivamente. Después, las respuestas d_x y d_y se ponderan con una función Gaussiana centrada en el punto de interés. Posteriormente, cada sub-región se describe por un vector $v = (\sum d_x, \sum |d_x|, \sum d_y, \sum |d_y|)$ y se concatenan los descriptores de todas las sub-regiones para formar el descriptor SURF asociado al punto característico. La Figura 3.2b muestra la cuadrícula de 4×4 sub-regiones asociadas a vectores descriptores que conforman SURF. El proceso descrito muestra que la diferencia entre píxeles es la base para formar el descriptor.

En el caso del descriptor SURF, la sumatoria de las derivadas en direcciones vertical y horizontal, y sus respectivos valores absolutos, aporta información de polaridad, es decir, de la dirección global de los gradientes, por lo cual, aunque no es en sí mismo un histograma de gradientes, se puede interpretar como una aproximación a uno de ellos a través de patrones de intensidad.

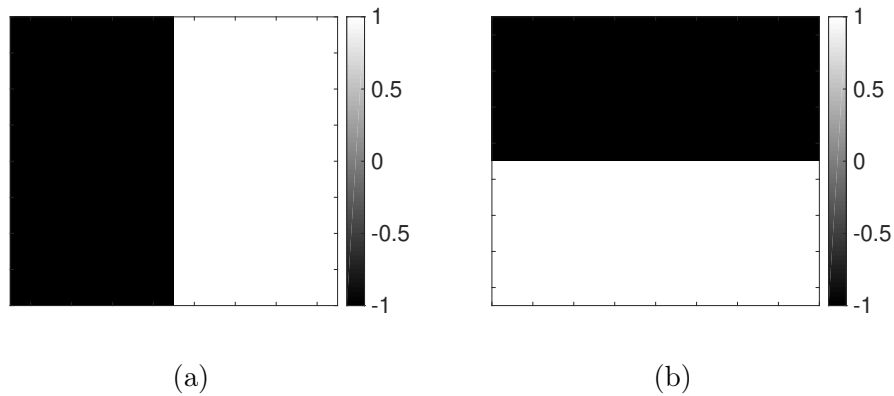


Figura 3.3: Filtros wavelet Haar usados en SURF
(a) horizontal. (b) vertical.

3.3 Evaluación del desempeño en el problema de correspondencia

La evaluación del desempeño en correspondencia entre imágenes se basa en métricas basadas en la repetibilidad de los puntos detectados en imágenes capturadas bajo diferentes condiciones, y en las verdaderas correspondencias en los grupos de descriptores dada una transformación geométrica conocida [Mikolajczyk and Schmid, 2005, Mikolajczyk et al., 2005].

3.3.1 Desempeño de los detectores

Los detectores de puntos de interés se evalúan con base en criterios de repetibilidad [Mikolajczyk et al., 2005]. La repetibilidad se obtiene al comparar el traslape entre las regiones detectadas en distintas imágenes. La Figura 3.4a ilustra la detección afín de un punto en el que la elipse azul representa la zona de influencia que debe incluirse en una posterior descripción. La Figura 3.4b ilustra la detección afín del mismo punto en la escena sobre una imagen distinta y la elipse amarilla representa la zona de influencia que debe incluirse en una posterior descripción. La Figura 3.4c muestra las dos zonas detectadas proyectadas al sistema de coordenadas de la Figura 3.4a.

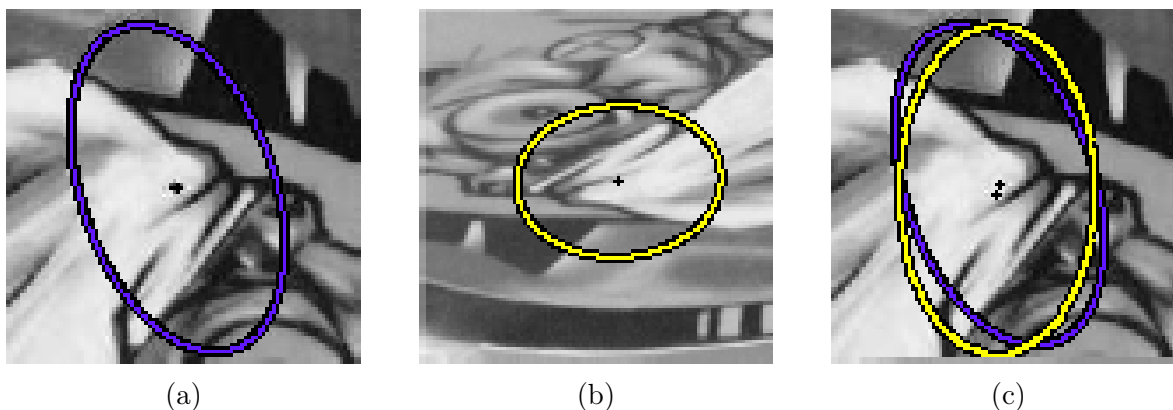


Figura 3.4: Detección en dos imágenes diferentes.

- (a) Un punto característico detectado en la imagen A. (b) Un punto característico detectado en la imagen B. (c) Detección de B proyectada en A. Adaptada de [Mikolajczyk et al., 2005].

El traslape entre las regiones se computa como la razón Intersección sobre Unión

(IoU) definida por

$$IoU = \frac{R_{\mu a} \cap R_{H^T \mu b H}}{R_{\mu a} \cup R_{H^T \mu b H}} \quad (3.1)$$

donde:

- $R_{\mu a}$ es la región elíptica de un punto en una imagen A.
- $R_{H^T \mu b H}$ es la región elíptica de un punto detectado en una imagen B, proyectada al sistema de coordenadas de la imagen A.
- H es la matriz homográfica que relaciona las imágenes A y B.

A partir de la razón IoU se estima el error de traslape $E = 1 - IoU$. Se considera que dos regiones corresponden entre sí cuando el error de traslape E entre ellas es menor a un umbral t_h . Finalmente, la repetibilidad se computa como la razón entre la cantidad de correspondencias detectadas y la menor cantidad de puntos detectados entre un par de imágenes.

Menores valores para el umbral t_h resultan en mayor precisión de las zonas detectadas que sean mutuamente correspondientes, como se ilustra en la Figura 3.5. Es de notar, que un detector ideal sería capaz de detectar las mismas regiones como correspondientes sin importar el umbral seleccionado, es decir, la variación del índice de repetibilidad respecto al cambio de umbral se asocia inversamente con la precisión del detector.

3.3.2 Desempeño de los descriptores

Los descriptores se evalúan utilizando un criterio basado en la cantidad de verdaderas y falsas correspondencias entre pares de imágenes. Las regiones A y B son identificadas como correspondientes si una métrica en el espacio de características, generalmente la distancia euclidiana, entre los respectivos descriptores D_A y D_B es

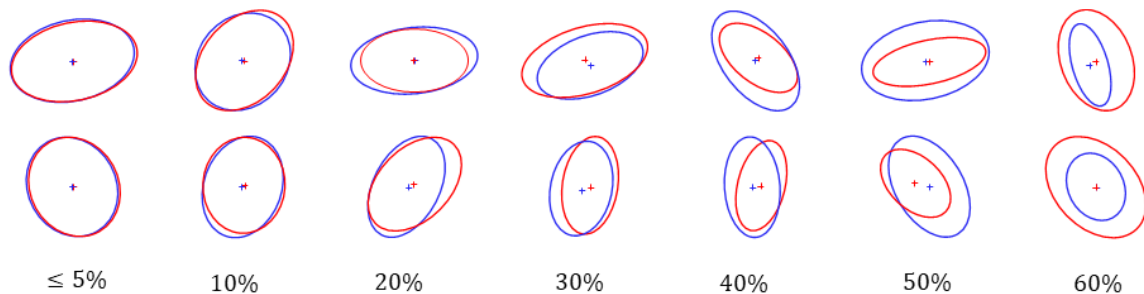


Figura 3.5: Errores de Traslape.
Tomada de [Mikolajczyk et al., 2005].

menor a un umbral t_h que para efectos de la evaluación varía. Cada descriptor de una imagen referencia se compara con todos los descriptores de la imagen con la que forma el par estereoscópico y, conocida la transformación geométrica entre las imágenes, se cuentan las verdaderas y falsas correspondencias. A partir de este conteo se estiman las métricas de desempeño *Razón de Verdaderas Correspondencias* (R) y *1-Precisión* ($1 - P$), definidas como

$$R = \frac{VC}{C} \quad (3.2)$$

y

$$1 - P = \frac{FC}{VC + FC} \quad (3.3)$$

donde:

- VC son las verdaderas correspondencias.
- FC son las falsas correspondencias.
- C son las posibles correspondencias.

Un descriptor con desempeño perfecto tendría la *Razón de verdaderas correspondencias* igual a uno para todos los valores de precisión. Curvas horizontales indican que la *Razón de Verdaderas Correspondencias* ha alcanzado la más alta precisión y se ve limitada por la especificidad de la escena. Una curva de lento crecimiento muestra que el descriptor se ve afectado por la degradación de las imágenes. La Figura 3.6 ilustra curvas típicas de R vs $1 - P$. La curva azul corresponde a un desempeño perfecto, la curva amarilla muestra un mejor desempeño

que la curva verde y la curva roja muestra un mejor desempeño que la curva magenta. Sin embargo, debido a la independencia entre la *Razón de Verdaderas Correspondencias* y la *Precisión* en el problema de correspondencia, las curvas que se obtienen de experimentos reales requieren de un método de comparación más sofisticado que la inspección visual.

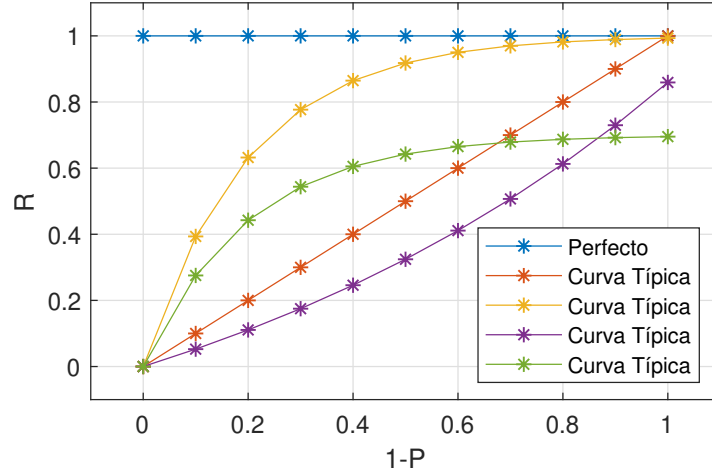


Figura 3.6: Curvas ideales de R vs $1 - P$.

Para comparar cuantitativamente el desempeño de los descriptores se utiliza la función F_β , definida como

$$F_\beta = (1 + \beta^2) \frac{P \cdot R}{\beta^2 \cdot P + R} \quad (3.4)$$

a partir de los datos de las gráficas R vs $1 - P$ con $\beta = 0.5$ de manera que se dé más peso a la *Precisión* que a la *Razón de verdaderas correspondencias* [Lee and Park, 2017].

3.4 Correspondencia

Las regiones A y B se identifican como correspondientes si una métrica en el espacio de características entre los respectivos descriptores D_A y D_B cumple con un conjunto de condiciones [Mikolajczyk and Schmid, 2005]. Las métricas más usadas son SSD (Sumatoria de diferencias al cuadrado) SAD (Sumatoria de diferencias absolutas). A

continuación se definen algunas estrategias de correspondencia:

- Basada en umbralización: A y B son correspondientes si, dado un umbral t_h , $\|D_A - D_B\| < t_h$. De esta manera, pueden identificarse múltiples puntos como correspondientes.
- Basada en el vecino más cercano: A y B son correspondientes si $\|D_A - D_B\| < \|D_A - D_{X \neq B}\| \wedge \|D_A - D_B\| < \|D_A - D_{X \neq B}\|$. De esta manera, puede identificarse solo un punto como correspondiente.
- Basada en proporción: A y B son correspondientes si $\frac{\|D_A - D_B\|}{\|D_A - D_C\|} < t_h$. Donde D_B es el vecino más cercano y D_C es el segundo vecino más cercano. De esta manera, puede identificarse solo un punto como correspondiente.

3.5 Color

El color es una característica de bajo nivel que brinda información relevante para diversas aplicaciones en visión por computador [Shamoi et al., 2014]. Un color puede representarse mediante su distribución espectral de potencia en un amplio rango de longitudes de onda, sin embargo, esta representación no guarda correspondencia con la forma en que los conos de la retina humana captan la luz en solo tres bandas espectrales (Rojo: 564 nm, Verde: 534 nm y Azul: 420 nm) [Ford and Roberts, 1998]. Debido a esto, se han propuesto diversos espacios de color [Alarcon and Dalmau, 2014, Alegre and Pajares, 2016], usualmente de tres dimensiones [Susstrunk et al., 1999]. Los espacios de color representan matemáticamente la reacción humana ante los estímulos en las tres longitudes de onda mencionadas [Zhou and Hansen, 2016]. La información contenida en las imágenes de color se relaciona directamente con la escena y se codifica de acuerdo con el espacio de color utilizado. Los espacios de color pueden clasificarse en Perceptuales y No Perceptuales [Shamoi et al., 2014, Cepeda-Negrete et al., 2017].

- **Espacios Perceptuales:** Buscan que los cambios de cromaticidad e intensidad percibidos por el sistema visual humano coincidan con las diferencias de la percepción del color mediante métricas. Entre estos se destacan los espacios

CIELAB, CIELUV y RLAB en los cuales la distancia euclidiana entre dos colores es proporcional a la diferencia percibida entre dichos colores [Cepeda-Negrete et al., 2017].

- **Espacios No Perceptuales:** No conservan uniformemente los cambios de cromaticidad ni intensidad de la percepción humana. Entre estos se destacan los espacios HSI, HSV, YIQ, YC_bC_r y $O_1O_2O_3$ [Cepeda-Negrete et al., 2017].

Los espacios de color HSV y HSI son ampliamente usados por su capacidad de desacople en la información de cromaticidad e intensidad, así como por la intuitividad de la representación cromática como matíz y saturación, sin embargo, los colores rojizos se separan y se presentan indeterminaciones en las componentes de cromaticidad para los grises. Para efectos de ilustración, se describen tres espacios No Perceptuales de color de amplio uso.

3.5.1 Espacio RGB

El espacio RGB es un estándar que se usa en la mayoría de dispositivos de captura y visualización de imágenes. Se basa en la adición de los iluminantes primarios Rojo (R), Verde (G) y Azul (B), generando una representación cartesiana de 3 dimensiones con origen en el color negro y cuyos vectores generadores corresponden a los iluminantes primarios. En este espacio, un color es descrito por una terna de valores, $C^{RGB} = (C^R, C^G, C^B) = \{C^b \in [0, 1] \mid b := \{R, G, B\}\}$, que representa la intensidad asociada a cada una de las bandas mencionadas sin ser lineal respecto a la percepción visual humana [Alegre and Pajares, 2016, Alonso Pérez, 2009]

La Figura 3.7 ilustra los colores obtenidos por los planos $R = 1$, $G = 0$ y $B = 1$. Las fronteras de este cubo son los colores pertenecientes a los planos $R = 0$, $G = 0$, $B = 0$, $R = 1$, $G = 1$ y $B = 1$.

En el espacio RGB, dos colores C_1 y C_2 se definen como $C_1^{RGB} = (C_1^R, C_1^G, C_1^B)$ y $C_2^{RGB} = (C_2^R, C_2^G, C_2^B)$. Puede considerarse que C_1 y C_2 tienen la misma cromaticidad con diferente intensidad si, en el espacio RGB, se cumple la condición de colinealidad

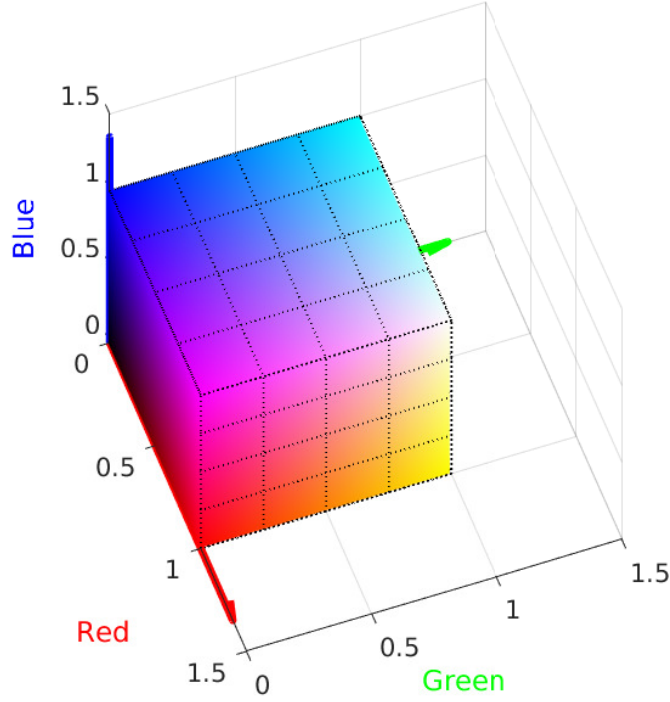


Figura 3.7: Distribución de colores en el espacio RGB.

definida por

$$C_2^{RGB} = \left(k \cdot C_1^R, \quad k \cdot C_1^G, \quad k \cdot C_1^B \right) = k \cdot C_1^{RGB} \quad (3.5)$$

$$\forall k \in \mathbb{R}^+ \setminus 1 \mid C_i^b \in [0, 1], b := \{R, G, B\}, i := \{1, 2\}$$

la cual indica que la intensidad del color tiene efectos sobre las 3 componentes de este espacio [Lynch et al., 2013]. Como las imágenes a color son capturadas por sensores en las bandas espectrales correspondientes a los colores primarios rojo, verde y azul, los ruidos en las mediciones sobre cada banda se propagan en el espacio RGB.

3.5.2 Espacios HSV, HSI y HSR

En el espacio HSI, que se inspira en el modo humano para describir los colores [Alegre and Pajares, 2016, Alonso Pérez, 2009], un color C es descrito mediante una terna de valores $C^{HSI} = (C^H, C^S, C^I)$. Inicialmente, se define C' como la intersección entre el segmento de línea \overline{OC} , comprendido entre el origen O y el color C , y el plano que cumple $R+G+B = 1$. Después, se define la componente C^H (Matiz) como el ángulo izquierdo medido sobre el plano $R + G + B = 1$ desde la línea $\overline{Gris - Rojo}$ hasta C' . La componente C^S (Saturación) corresponde a la proporción de pureza de C' frente a su matiz predominante. Finalmente, la componente C_I corresponde a la intensidad de C medida como la proyección de C sobre la línea de grises. La transformación $RGB \rightarrow HSI$ se realiza mediante

$$\begin{bmatrix} C^H \\ C^S \\ C^I \end{bmatrix} = \begin{bmatrix} atan2(\sqrt{3}(C^G - C^B), 2C^R - C^G - C^B) \\ 1 - \frac{3 \min(C^R, C^G, C^B)}{C^R + C^G + C^B} \\ \frac{C^R + C^G + C^B}{3} \end{bmatrix} \quad (3.6)$$

donde C^R , C^G y C^B son las componentes que describen al color C en el espacio RGB. La Figura 3.8 ilustra la relación de los componentes entre los espacios HSI y RGB.

Los espacios HSI y HSV tienen las mismas componentes de cromaticidad pero difieren en la componente de intensidad. La transformación $RGB \rightarrow HSV$ se realiza mediante la expresión definida en:

$$\begin{bmatrix} C^H \\ C^S \\ C^V \end{bmatrix} = \begin{bmatrix} atan2(\sqrt{3}(C^G - C^B), 2C^R - C^G - C^B) \\ 1 - \frac{3 \min(C^R, C^G, C^B)}{C^R + C^G + C^B} \\ \max(C^R, C^G, C^B) \end{bmatrix} \quad (3.7)$$

donde C^R , C^G y C^B son las componentes que describen al color C en el espacio RGB.

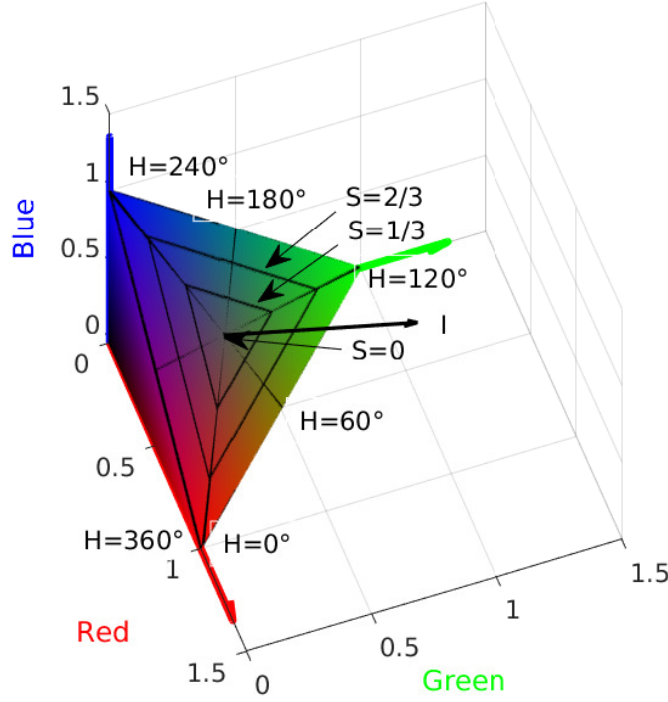


Figura 3.8: Distribución de colores en el espacio HSI

Las representaciones en los espacios HSI y HSV, obtenidas mediante las Ecuaciones 3.6 y 3.7, de un color C_2 que tiene la misma cromaticidad de C_1 con distinta intensidad, es decir que cumple la condición descrita en la Ecuación 3.5, son $(C_2^H, C_2^S, C_2^I) = (C_1^H, C_1^S, k \cdot C_1^I)$ en el espacio HSI, y $(C_2^H, C_2^S, C_2^V) = (C_1^H, C_1^S, k \cdot C_1^V)$ en el espacio HSV. Estas representaciones muestran que las componentes (C^H, C^S) describen la cromaticidad de manera desacoplada a las componentes de intensidad C_I y C_V .

Adicionalmente, los espacios HSV y HSI se han representado utilizando las componentes de un sistema de coordenadas esféricas cuyo eje polar coincide con la línea de grises [Chen et al., 2013b]. Este espacio será llamado Hue Saturation and Ratio (Versión en coordenadas esféricas de HSI) (HSR) en este documento. En HSR, un color C se representa mediante la terna $(C^\rho, C^\theta, C^\varphi)$, donde C^θ conserva la interpretación de Matiz, C^φ se interpreta como la Saturación sin la normalización ante el matiz predominante y C^ρ se interpreta como una forma de intensidad. La

transformación entre el espacio RGB y la forma esférica de HSI se define en:

$$\begin{bmatrix} C^\rho \\ C^\theta \\ C^\varphi \end{bmatrix} = \begin{bmatrix} \sqrt{C^R{}^2 + C^G{}^2 + C^B{}^2} \\ \text{atan2}(\sqrt{3}(C^G - C^B), 2C^R - C^G - C^B) \\ \arccos \frac{C^R + C^G + C^B}{\sqrt{3(C^R{}^2 + C^G{}^2 + C^B{}^2)}} \end{bmatrix} \quad (3.8)$$

donde C^R , C^G y C^B son las componentes que describen al color C en el espacio RGB.

La representación esférica de los espacios HSI y HSV, obtenida mediante las Ecuaciones 3.8, de un color C_2 que tiene la misma cromaticidad de C_1 con distinta intensidad, es decir que cumple la condición descrita en la Ecuación 3.5, es $(C_2^\rho, C_2^\theta, C_2^\varphi) = (k \cdot C_1^\rho, C_1^\theta, C_1^\varphi)$. Esto muestra que las componentes (C^θ, C^φ) describen la cromaticidad de manera desacoplada a la componente de intensidad C^ρ .

3.5.3 Colores oponentes

En este espacio, un color C se describe mediante una terna de valores $C^{O_1 O_2 O_3} = (C^{O_1}, C^{O_2}, C^{O_3})$. Este espacio corresponde a una rotación del espacio RGB en el que uno de los vectores generadores coincide con la línea de grises. La transformación $RGB \rightarrow O_1 O_2 O_3$ se realiza mediante

$$\begin{bmatrix} C^{O_1} \\ C^{O_2} \\ C^{O_3} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} C^R \\ C^G \\ C^B \end{bmatrix} \quad (3.9)$$

donde C^R , C^G y C^B son las componentes que describen al color C en el espacio RGB.

Se resalta que la componente C^{O_3} de este espacio es proporcional a la componente C^I de HSI, por lo que se considera que contiene la información de intensidad mientras

las componentes (C^{O1}, C^{O2}) contienen la información cromática. Sin embargo, dada la representación en colores oponentes $O_1O_2O_3$, obtenida mediante la Ecuación 3.9, de un color C_2 que tiene la misma cromaticidad de C_1 con distinta intensidad, es decir que cumple la condición descrita en la Ecuación 3.5, es $(C_2^{O1}, C_2^{O2}, C_2^{O3}) = (k \cdot C_1^{O1}, k \cdot C_1^{O2}, k \cdot C_1^{O3})$, se observa que la intensidad tiene efectos sobre las 3 componentes de este espacio, por lo cual este espacio no desacopla la información cromática de la intensidad.

3.5.4 Funciones invariantes a color

Algunas funciones resultan invariantes a los cambios cromáticos al eliminar la dependencia de la intensidad en la obtención de una única capa de información con contenido cromático. En el problema de correspondencia se ha propuesto la función definida por [Abdel-Hakim and Farag, 2006]

$$C^C = \frac{E_\lambda}{E_{\lambda\lambda}} \quad (3.10)$$

en la que E_λ y $E_{\lambda\lambda}$ representan coeficientes diferenciales espectrales y se aproximan linealmente como

$$\begin{bmatrix} E \\ E_\lambda \\ E_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.3 & 0.04 & -0.35 \\ 0.34 & -0.6 & 0.17 \end{bmatrix} \cdot \begin{bmatrix} C^R \\ C^G \\ C^B \end{bmatrix} \quad (3.11)$$

donde C^R , C^G y C^B son las componentes que describen al color C en el espacio RGB. Así, dada la representación en función invariante a color C , obtenida mediante la Ecuación 3.10, de un color C_2 que tiene la misma cromaticidad de C_1 con distinta intensidad, es decir que cumple la condición descrita en la Ecuación 3.5, es $C_2^C = C_1^C$, se observa que la intensidad del color no tiene efecto sobre esta función.

3.6 Discusión

- El uso de diferencias entre valores de píxeles dentro de la vecindad de los puntos característicos disminuye la sensibilidad de la descripción respecto a alteraciones fotométricas en las imágenes. Por otra parte, la geometría con que se agrupa la información permite la obtención de descriptores que contienen la distribución de las diferencias respecto a la orientación y ubicación relativa de cada zona. Debido a lo anterior, el desempeño de los descriptores basados en histogramas de gradientes depende también de las geometrías y las representaciones de la información cromática y de intensidad que son fuente en la estimación de gradientes.
- Aunque los espacios de color HSI y HSV son ampliamente usados por su capacidad de desacoplar la información de cromaticidad de la información de intensidad, la discontinuidad presente para los tonos rojizos introduce errores en la estimación de diferencias cromáticas, y estos errores afectan la estabilidad de los descriptores locales que se estiman a partir de sus componentes cromáticas.

4. Propuesta de un espacio de color para los descriptores basados en histogramas de gradientes

En los descriptores basados en histogramas de gradientes contruidos mediante la concatenación de descriptores sobre componentes de un espacio de color, se usan aproximaciones a las derivadas a partir de diferencias entre píxeles. Esto supone que todas las componentes forman un espacio euclídeo, lo que es verdadero para el espacio RGB y de colores oponentes $O_1O_2O_3$, pero en estos espacios de color no se desacopla la información de cromaticidad de la de intensidad. Por otra parte, el supuesto es falso para espacios como el HSI y HSV y además, dadas sus unidades y los rangos de las componentes H (un ángulo) y S (una proporción), se introducen errores al encontrar distancias cromáticas de forma euclidiana. Para mejorar el desempeño de los descriptores basados en histogramas de gradientes mediante una representación de color, se propone un espacio de color no perceptual basado en coordenadas esféricas llamado RTP (Rho (ρ), Theta (θ), Phi (φ)) que desacopla la cromaticidad (C^θ, C^φ) de la intensidad (c_ρ), representa de manera continua las cromaticidades y unifica las unidades en la información de cromaticidad. El uso de este espacio permitió reducir sustancialmente el error en la estimación de distancias cromáticas aunque sin replicar la sensibilidad humana.

4.1 Definición del espacio de color RTP

Un sistema de coordenadas esférico queda completamente determinado por un origen O , un círculo fundamental CF , un punto fundamental PF y los sentidos de medición de los ángulos [Berrocoso et al., 2003]. A continuación se desarrollan las expresiones de transformación $RGB \leftrightarrow RTP$.

Tomando como base el espacio euclídeo de tres dimensiones RGB (Sección 3.5), en el que $x = (x^R, x^G, x^B)$ describe un color determinado, se define:

- El plano π_1 que contiene la línea de grises ($x^R = x^G = x^B$) y el color azul $(0, 0, 1)$:

$$\pi_1 = \{x \in \mathbb{R}^3 \mid x = s \cdot (0, 0, 1)^T + t \cdot (1, 1, 1)^T \forall s, t \in \mathbb{R}\} \quad (4.1)$$

- El plano π_2 que es perpendicular a π_1 y contiene la línea de grises ($R = G = B$):

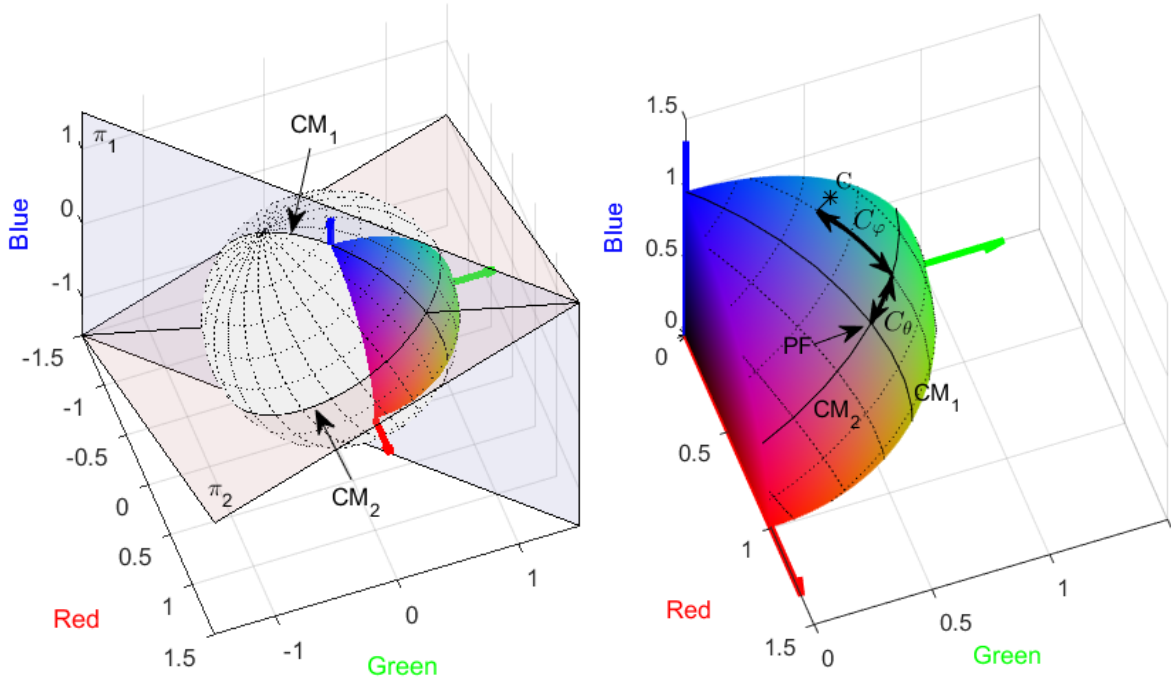
$$\pi_2 = \{x \in \mathbb{R}^3 \mid x = s \cdot [(0, 0, 1) \times (1, 1, 1)]^T + t \cdot (1, 1, 1)^T \forall s, t \in \mathbb{R}\} \quad (4.2)$$

- La esfera de radio unitario E cuyo centro se ubica en el color negro $(0, 0, 0)$:

$$E = \{x \in \mathbb{R}^3 \mid \|\vec{x}\| = 1\} \quad (4.3)$$

Las intersecciones de la esfera E (Ecuación 4.3) con los planos π_1 (Ecuación 4.1) y π_2 (Ecuación 4.2) definen los círculos máximos CM_1 y CM_2 , respectivamente. El color negro se toma como origen, el círculo CM_2 se toma como círculo fundamental y la intersección entre CM_1 y CM_2 en el octante $x^b \in [0, 1] \mid b := \{R, G, B\}$ se toma como punto fundamental PF . A partir de estos elementos se determinan las componentes del sistema de coordenadas esféricas para el espacio de color RTP (ver Figura 4.1a). En el espacio de color RTP, un color C se representa como una terna de valores en un sistema de coordenadas esféricas $C^{RTP} = (C_\rho, C_\theta, C_\varphi)$ con origen O en el color negro. La componente C_ρ es la norma de \vec{OC} . La componente C_θ es el ángulo que forma \vec{OC} respecto al plano π_1 , tomando como dirección positiva del color rojo al verde. La componente C_φ es el ángulo que forma \vec{OC} respecto al plano π_2 , tomando como dirección positiva del color amarillo al azul. Las componentes angulares (C_θ, C_φ) corresponden a los arcos medidos sobre la esfera unitaria E medidos a partir de los círculos máximos CM_1 y CM_2 , como se ilustra en la Figura 4.1b.

Dado un color C que es representado en el espacio RGB como $C^{RGB} = (C^R, C^G, C^B)$, se puede encontrar su representación C^{RTP} en el espacio RTP



(a) Geometría del espacio RTP.

(b) Distribución de colores en el espacio RTP.

Figura 4.1: Definición geométrica del espacio RTP.

utilizando la transformación dada por

$$C^{RTP} = \begin{bmatrix} C^\rho \\ C^\theta \\ C^\varphi \end{bmatrix} = \begin{bmatrix} \sqrt{C^R{}^2 + C^G{}^2 + C^B{}^2} \\ \arctan \left(-\sqrt{\frac{3}{2}} \frac{C^R - C^G}{C^R + C^G + C^B} \right) \\ \arctan \left(\frac{-C^R - C^G + 2C^B}{\sqrt{5C^R{}^2 + 5C^G{}^2 + 2C^B{}^2 + 4C^B C^R + 4C^G C^B + 2C^R C^G}} \right) \end{bmatrix} \quad (4.4)$$

La transformación contraria, es decir, encontrar la representación en el espacio RGB







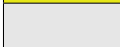
a partir de la representación en el espacio RTP, está dada por

$$C^{RGB} = \begin{bmatrix} C^R \\ C^G \\ C^B \end{bmatrix} = C^\rho \cdot \begin{bmatrix} \frac{\sqrt{3}}{3} \cos C^\varphi \cos C^\theta - \frac{\sqrt{2}}{2} \cos C^\varphi \sin C^\theta - \frac{\sqrt{6}}{6} \sin C^\varphi \\ \frac{\sqrt{3}}{3} \cos C^\varphi \cos C^\theta + \frac{\sqrt{2}}{2} \cos C^\varphi \sin C^\theta - \frac{\sqrt{6}}{6} \sin C^\varphi \\ \frac{\sqrt{3}}{3} \cos C^\varphi \cos C^\theta - \sqrt{\frac{2}{3}} \sin C^\varphi \end{bmatrix} \quad (4.5)$$

Los colores C_1 y C_2 que cumplen con la condición de colinealidad de la Ecuación 3.5 se representan en el espacio RTP como $C_1^{RTP} = (C_1^\rho, C_1^\theta, C_1^\varphi)$ y $C_2^{RTP} = (k \cdot C_1^\rho, C_1^\theta, C_1^\varphi)$. Así, en el espacio de color RTP, las componentes angulares (C^θ, C^φ) describen la cromaticidad de manera desacoplada a la componente de intensidad C^ρ .

4.2 Distancias cromática HS vs. θ φ

Las imágenes a color son capturadas por sensores en el espacio RGB (Ver sección 3.5.1) y el ruido de las mediciones se propaga hacia los espacios en los que se realice el procesamiento, generando distintas distribuciones de acuerdo con la ubicación, e incertidumbre al estimar diferencias y gradientes cromáticos. El Cuadro 4.1 muestra un listado de 7 colores definidos con sus respectivas representaciones en los espacios de color RGB, HSI y RTP cuyas componentes cromáticas, en los espacios HSI y RTP, se ilustran en las Figuras 4.2a y 4.2b que muestran la distribución de colores con intensidad constante en los espacios de color HSI y RTP, respectivamente.

Color	RGB			HSI			RTP		
	R	G	B	H [rad]	S	I	ρ	θ [rad]	φ [rad]
	0.90	0.09	0.09	0.0016	0.7500	0.36	0.909	-0.7430	-0.3723
	0.09	0.90	0.09	2.0944	0.7500	0.36	0.909	0.7430	-0.3723
	0.09	0.09	0.90	4.1888	0.7500	0.36	0.909	0.0000	-0.8148
	0.09	0.90	0.90	3.1400	0.8571	0.63	1.276	0.4834	0.2622
	0.90	0.09	0.90	5.2360	0.8571	0.63	1.276	-0.4834	0.2622
	0.90	0.90	0.09	1.0472	0.8571	0.63	1.276	0.0000	-0.5449
	0.90	0.90	0.90	1.5708	0.0000	0.90	1.559	0.0000	0.0000

Cuadro 4.1: 7 Colores representados en tres espacios de color.

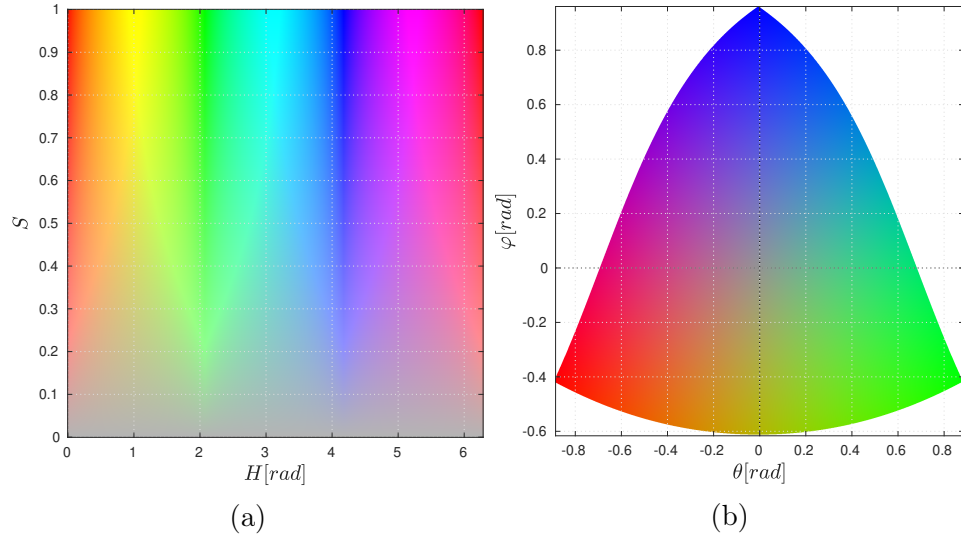
















Figura 4.2: Gamut de los espacios (a) RTP y (b) HSI/HSV.

En el espacio RGB, el ángulo entre dos vectores que comienzan en el origen y terminan en dos colores determinados, puede ser interpretado como un indicador de diferencia cromática [Gómez, 2016, Lynch et al., 2013]. El Cuadro 4.2 lista los ángulos, en radianes, estimados entre los colores listados en el Cuadro 4.1.

							
	0	—	—	—	—	—	—
	1.4859	0	—	—	—	—	—
	1.4005	1.4005	0	—	—	—	—
	1.3808	0.6856	0.7342	0	—	—	—
	0.6856	1.3808	0.7342	0.9664	0	—	—
	0.7627	0.7627	1.3597	0.9407	0.9407	0	—
	0.8311	0.8311	0.8148	0.5499	0.5499	0.5449	0

Cuadro 4.4: Distancias cromáticas calculadas en el espacio RTP (θ, φ) .

El símbolo — indica que, dada la simetría, el valor puede encontrarse en la sección inferior.

La Figura 4.3 ilustra la relación de las distancias cromáticas calculadas como distancias euclidianas sobre las componentes (H, S) , de los espacios HSI y HSV, y (θ, φ) , del espacio RTP, con el ángulo entre colores calculado en el espacio RGB. Esta gráfica muestra que hay una relación aproximadamente lineal entre las diferencias cromáticas y la distancia euclidiana encontradas con las componentes cromáticas del espacio RTP. Por otra parte, la distancia euclidiana encontrada con las componentes cromáticas de los espacios HSI y HSV no es consistente con la distancia cromática estimada mediante ángulos, lo cual evidencia que los espacios de color HSI y HSV no son compatibles con los supuestos euclidianos sobre los que se basa el cálculo de gradientes.

Con base en los análisis anteriores respecto al espacio de color propuesto, se puede decir que:

- La componente radial C^ρ contiene la información de intensidad, dejando la información de cromaticidad contenida en las componentes angulares (C^θ, C^φ) .
- Las componentes angulares (C^θ, C^φ) del espacio propuesto quedan indeterminadas para el color negro, sin embargo, este queda completamente definido con la componente radial $C^\rho = 0$.
- La recta $(C^\rho, 0, 0)$ corresponde a la recta de tonos de gris.

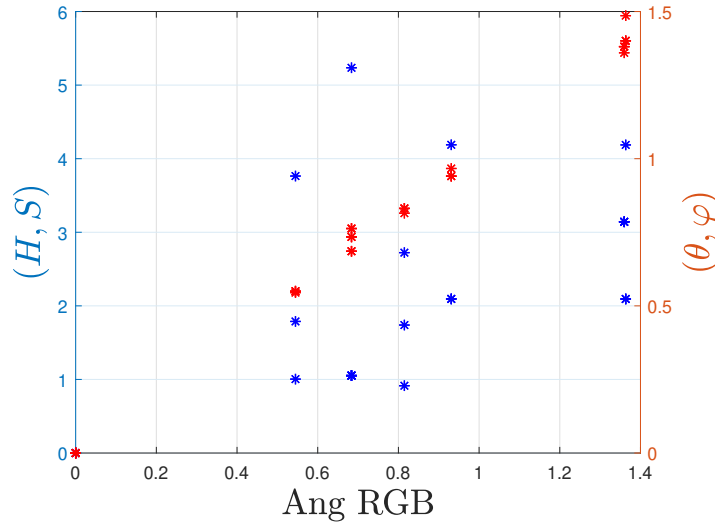


Figura 4.3: Relación entre de distancias cromáticas en los espacios de color RGB, HSI/HSV y RTP.

- La trayectoria de radio normalizado $(1, 0, C^\varphi)$ muestra el cambio de amarillo a azul pasando por el gris con una intensidad normalizada de 1.
- La trayectoria de radio normalizado $(1, C^\theta, \text{mín } C^\varphi)$ muestra el cambio de rojo a verde pasando por el amarillo con una intensidad normalizada de 1.

Finalmente, la Figura 4.4a muestra una imagen sintética que exhibe un amplio rango de cromaticidades conocida como Círculo Cromático. La representación en los espacios de color RGB, HSV, HSI y RTP se ilustra en la Figura 4.4b, donde se puede observar que existe una relación directa entre la componente I del espacio HSI y la componente ρ del espacio RTP. Las componentes (H, S) de los espacios HSV y HSI codifican la cromaticidad de la imagen de la misma manera y presentan un cambio abrupto en la zona correspondiente a los rojos aunque las cromaticidades son cercanas. Este cambio no se presenta en las componentes (θ, φ) del espacio RTP, lo cual constituye una ventaja en la descripción del color, pues las diferencias entre colores siempre son consistentes.

4.3 Discusión

- El espacio de color propuesto (RTP) es no perceptual y se basa en observaciones de estudios previos según los cuales el color se puede representar de manera

satisfactoria con coordenadas esféricas. Como ocurre con los espacios HSI y HSV, RTP desacopla la información de cromaticidad de la información de intensidad, pero no introduce discontinuidades en los colores correspondientes a tonalidades de rojo. Además, en el espacio RTP se pueden encontrar diferencias cromáticas entre colores mediante aproximaciones euclidianas sin introducir errores debido a discontinuidades.

- Dada la geometría del espacio RTP, se espera que al usar sus componentes se mejore la robustez de los detectores ante cambios fotométricos locales, basada en MSER, y de descripción, basada en histogramas de gradientes como SIFT y SURF, y que esta mejora se aprecie con las métricas de *Razón de Verdaderas Correspondencias* y *Precisión*. El efecto del espacio RTP sobre las etapas de detección y descripción se verifica en la sección 5.

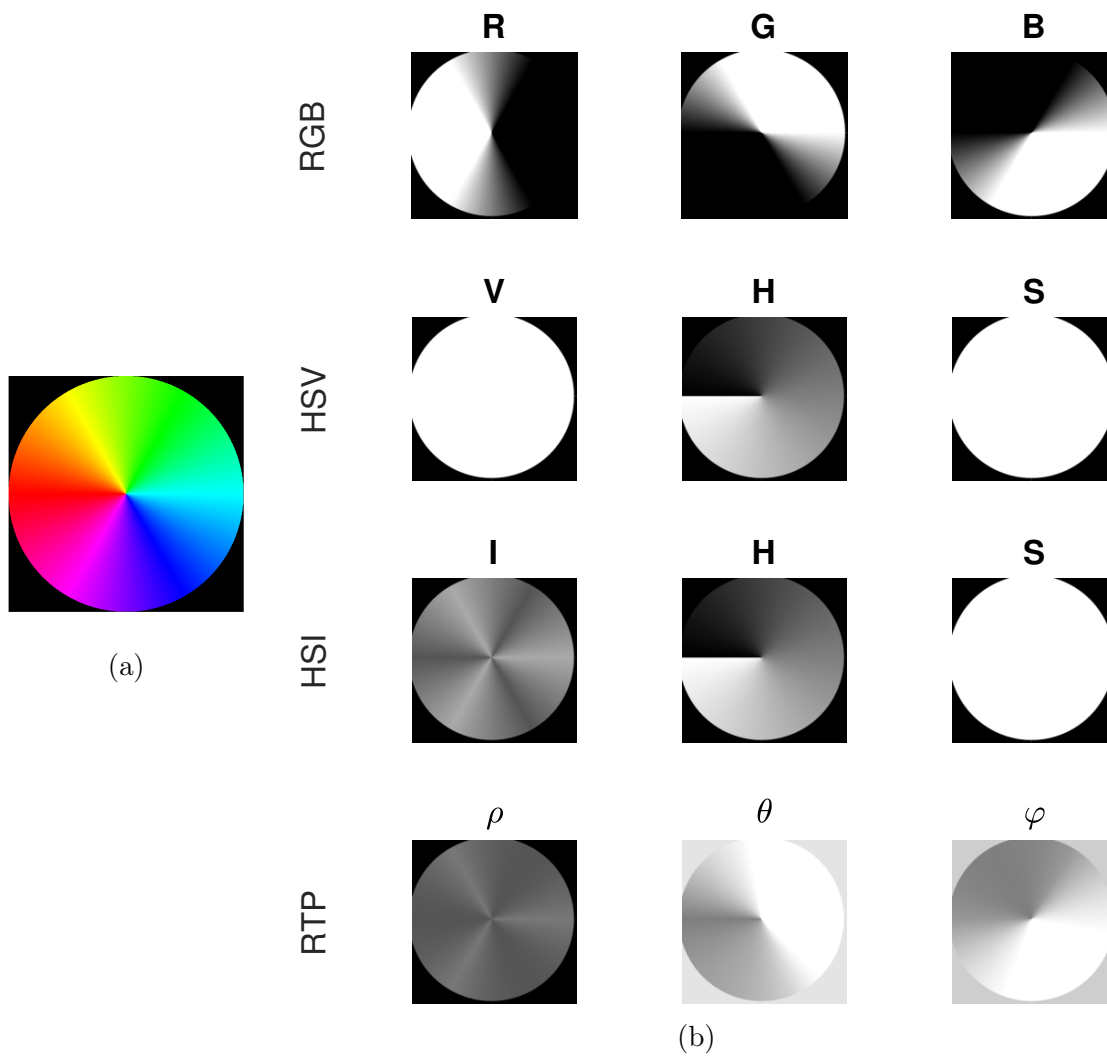


Figura 4.4: (a) Circulo cromático. (b) Descomposiciones del circulo cromático en los espacios RGB, HSV, HSI y RTP.

5. Evaluación de los aportes del espacio de color RTP sobre la correspondencia entre imágenes

Para evaluar el efecto del espacio de color en el problema de puesta en correspondencia, se adopta un método semiglobal para dos imágenes en color (Ver Figura 5.1) constituido por cuatro etapas. En la primera etapa se realiza la transformación a un espacio de color X . En la segunda etapa se detectan los puntos característicos sobre la componente de intensidad del espacio X . En la tercera etapa se emplea un descriptor basado en histogramas de gradientes sobre cada componente del espacio X y se concatenan los respectivos vectores de características para formar un descriptor con información de color. En la cuarta etapa se emplea la métrica de distancia euclidiana en el espacio de características usando como criterio de correspondencia el vecino más cercano.

Se realiza una evaluación comparativa del desempeño de la correspondencia al usar como fuente de información las imágenes representadas en tonos de gris, la función invariante a color C [Abdel-Hakim and Farag, 2006], los espacios de color RGB, colores oponentes $O_1O_2O_3$, HSI, HSR y el espacio RTP propuesto.

La evaluación se realiza con base en el desempeño alcanzado en las etapas de detección (etapa 2) y descripción (etapa 3) con cada transformación y espacio mencionado. Para ello se emplean las métricas de *Repetibilidad* (Sección 3.3.1) en la etapa 2, y *Razón de verdaderas correspondencias*, *Precisión* y $F_{0.5}$ en la etapa 3 (Sección 3.3.2). Las pruebas se realizan con imágenes reales de distintas escenas que están afectadas por transformaciones fotométricas y geométricas que pueden modelarse como matrices de homografía. La puesta en correspondencia con el espacio RTP se valida construyendo un mosaico con imágenes aéreas de una escena de geometría plana.

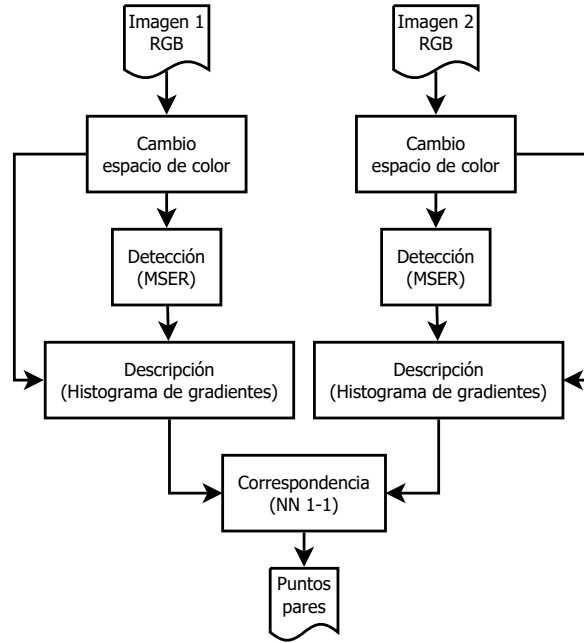


Figura 5.1: Esquema de la solución propuesta.

5.1 Data sets

Las pruebas se realizan con dos *data sets* de imágenes. El primero corresponde a un *data set* de dominio público, [Oxford University, 2004], constituido por 7 conjuntos de imágenes reales de distintas escenas que reproducen las transformaciones geométricas y fotométricas propias del problema de puesta en correspondencia [Mikolajczyk and Schmid, 2005, Mikolajczyk et al., 2005]. El segundo *data set* está constituido por un conjunto de imágenes aéreas del campus Meléndez de la Universidad del Valle que tienen las transformaciones típicas de un vuelo fotogramétrico [García Moreno et al., 2019].

Las imágenes que conforman los *data sets* son de escenas planas o fueron capturadas con cámara en posición fija. De esta manera se garantiza que la relación geométrica entre pares de imágenes se modela como una transformación homográfica. Cada conjunto consta de 6 imágenes de las que se toma una como referencia y con el resto se conforman los pares estereoscópicos $1 \leftrightarrow X | X \in \{2, 3, 4, 5, 6\}$ cuya matriz homográfica H hace parte del *data set*.

La Figura 5.2 muestra algunas de las imágenes que conforman los *data sets*. Las

Figuras 5.2a (Conjunto *Wall*) y 5.2b (Conjunto *Graffiti*) hacen parte de los conjuntos de imágenes con transformaciones debidas a cambios en los puntos de vista. Las Figuras 5.2c (Conjunto *Bikes*) y 5.2d (Conjunto *Trees*) hacen parte de los conjuntos de imágenes con transformaciones debidas al desenfoque de la cámara. Las Figuras 5.2e (Conjunto *Bark*) hacen parte del conjunto de imágenes con transformaciones debidas a rotaciones y cambios de escala. Las Figuras 5.2f (Conjunto *Leuven*) hacen parte del conjunto de imágenes con transformaciones debidas a cambios en la iluminación de la escena. Las Figuras 5.2g (Conjunto *UBC*) hacen parte de los conjuntos de imágenes con transformaciones debidas a la compresión. Las Figuras 5.2h (Conjunto *Univalle*) hacen parte de las imágenes aéreas con grandes cambios de punto de vista en condiciones normales de un vuelo fotogramétrico.

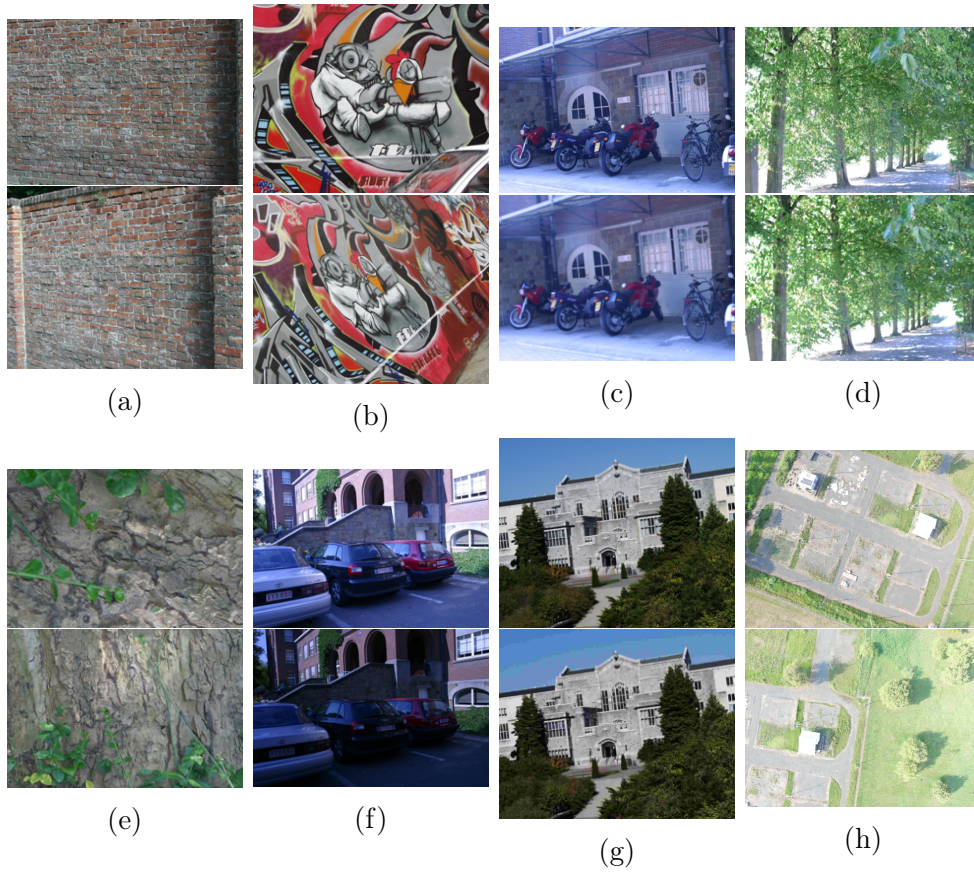


Figura 5.2: Ejemplo de imágenes del *data set*.

Las imágenes permiten analizar el efecto de las variaciones ante cambios en (a) y (b) puntos de vista, (c) y (d) enfoque, (e) rotación y escala, (f) iluminación, (g) compresión y (h) efectos en las imágenes aéreas.

5.2 Evaluación del detector MSER

Para evaluar el aporte del espacio de color RTP sobre la etapa de detección de puntos característicos basada en MSER se utiliza el protocolo de pruebas estándar descrito en la sección 3.3.1 con las imágenes mencionadas en la sección 5.1. Con cada par estereoscópico del *data set*, representado con las componentes de intensidad ρ del espacio RTP e I del espacio HSI, se varía el umbral t_h (Ver sección 3.3.1) entre 10 % y 60 % con pasos de 10 % y se generan las curvas de repetibilidad $R(E)$ en función del error de traslape. Después, el umbral de error de traslape se fija en el 40 % para comparar el desempeño del detector MSER en la detección de regiones de interés a partir de las componentes ρ e I . Se resalta que la componente I es proporcional la componente de intensidad del espacio de colores oponentes $O_1O_2O_3$. En el Apéndice A se registran las curvas de desempeño en detección en cada uno de los conjuntos evaluados. Una inspección visual de las curvas revela que la detección MSER con las componentes I y ρ tienen comportamientos similares.

Transformación	Conjunto	min(R40 %)		max(R40 %)		prom(R40 %)	
		ρ	I	ρ	I	ρ	I
Puntos de vista	Graf	49.25	48.49	80.6	79.32	65.78	64.48
	Wall	33.62	34.07	70.95	71.07	54.75	55.13
Rotación y escala	Bark	26.32	28.95	69.39	69.07	49.07	48.74
Enfoque	Trees	31.62	32.5	45.95	46.61	39.07	39.62
	Bikes	39.53	39.24	73.29	74.04	55.22	57.12
Compresión	UBC	26.83	28.91	65.33	63.50	50.31	50.01
Iluminación	Leuven	73.62	74.38	81.03	79.90	77.93	77.85
Aéreas	Univale	9.79	9.78	22.73	22.88	15.02	15.10

Cuadro 5.1: Resumen de los valores de repetibilidad con umbral fijo del 40 %.

El Cuadro 5.1 lista los valores mínimos, máximos y promedios obtenidos en la repetibilidad con un umbral de error de traslape del 40 % para cada conjunto de imágenes evaluado. En cada caso, se han resaltado los mejores valores de repetibilidad entre los correspondientes a los espacios de color RTP (50.89 % en promedio) y HSI (51.01 % en promedio). Puede observarse que la mayor diferencia en la *Repetibilidad*

de la detección sobre las componentes de intensidad I y ρ se obtiene con el conjunto *Bikes*, cuyas curvas de desempeño se ilustran en la Figura 5.3.

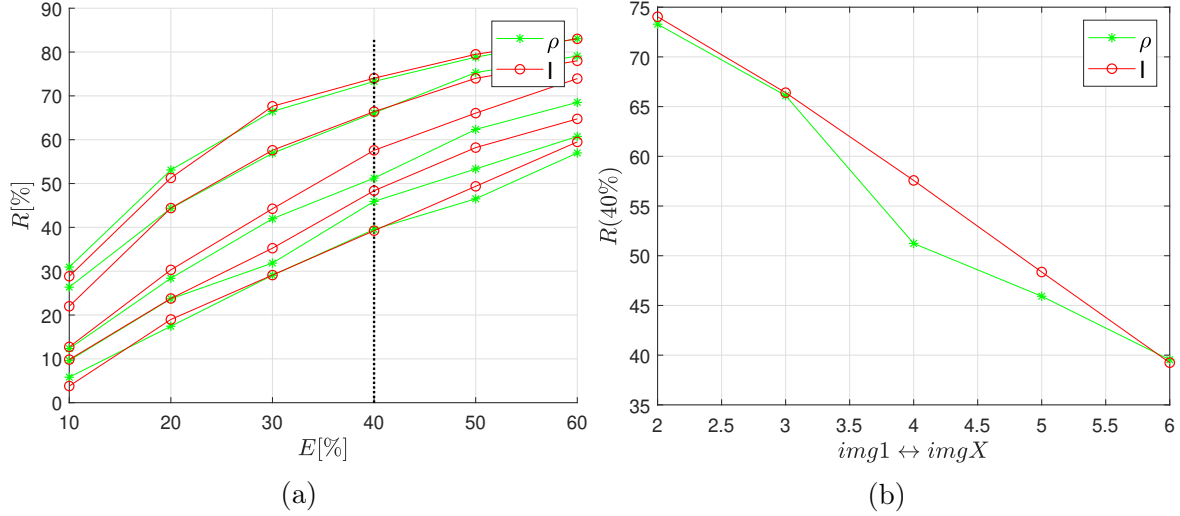


Figura 5.3: Desempeño en detección ante cambios de enfoque (Conjunto Bikes).
(a) *Repetibilidad* vs *Error de traslape* para cada par estereoscópico. (b) *Repetibilidad* vs *Par estereoscópico* con un *Error de traslape* fijo del 40 %.

La Figura 5.3a ilustra las curvas de *Repetibilidad* obtenidas al usar los pares estereoscópicos $1 \leftrightarrow X$ del conjunto de conjunto *Bikes* ($X \in \{2, 3, 4, 5, 6\}$). Se observa que la *Repetibilidad* aumenta para las componentes I y ρ junto con el error de traslape para el conteo de puntos correspondientes y disminuye con el nivel de distorsión de la imagen X con que se forma el par $1 \leftrightarrow X$. La Figura 5.3b muestra las curvas de *Repetibilidad* que se obtienen al fijar un error de traslape de 40 % para los pares estereoscópicos $1 \leftrightarrow X$ del conjunto de conjunto *Bikes*. Se observa que la repetibilidad obtenida con la componente ρ del espacio RTP (55.22 % en promedio) es menor en todas las variaciones de enfoque de escena *Bikes* que la asociada con la componente I del espacio HSI (57.12 % en promedio). Además, se observa que el par conformado por la imagen de referencia y la imagen número 4 del conjunto *Bikes* presenta la diferencia más alta (mayor al 5 %) en *Repetibilidad* con un error de traslape del 40 %.

La Figura 5.4 ilustra las regiones detectadas en la imagen número 4 del conjunto *Bikes* utilizando el detector MSER. Encerradas por elipses verdes de trazo continuo, se resaltan las regiones detectadas sobre la componente de intensidad del espacio RTP. Las

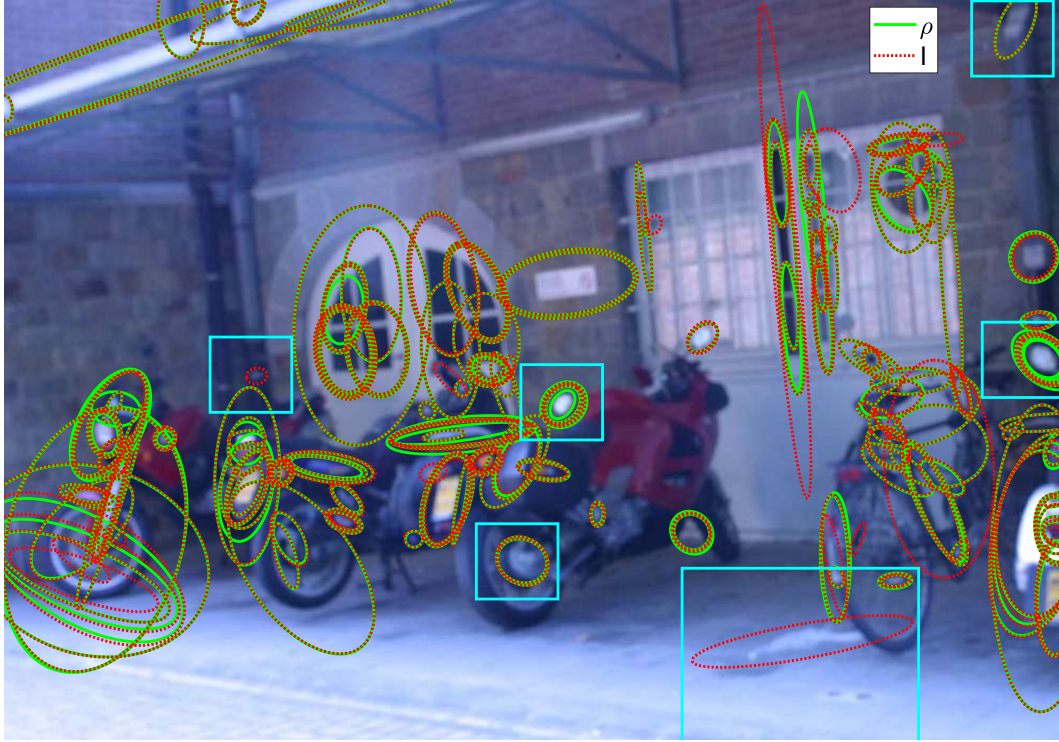


Figura 5.4: Regiones detectadas en la imagen número 4 del conjunto *Bikes*

elipses rojas con trazo punteado resaltan las regiones detectadas sobre la componente de intensidad del espacio HSI. Los recuadros de color cyan resaltan algunas regiones que se amplían en la Figura 5.5 y se describen a continuación.

Las Figuras 5.5a y 5.5d ilustran dos regiones que fueron detectadas de la misma manera con MSER sobre las componentes de intensidad I y ρ . La región de la Figura 5.5a fue detectada con dos niveles de umbral (ver sección 3.1), por lo que se visualizan dos áreas concéntricas pero que coinciden completamente entre las componentes I y ρ . Por su parte, la región ilustrada en la Figura 5.5d fue detectada con umbrales únicos y equivalentes entre las componentes I y ρ .

Las Figuras 5.5b y 5.5e ilustra dos regiones que fueron detectadas únicamente con la componente I . Este efecto aparece en regiones de poco contraste cromático que se ven más afectadas por el desenfoque en el espacio RTP para las zonas grises rodeadas de zonas también grises.

Las Figuras 5.5c y 5.5f ilustra dos regiones detectadas con más umbrales en el espacio

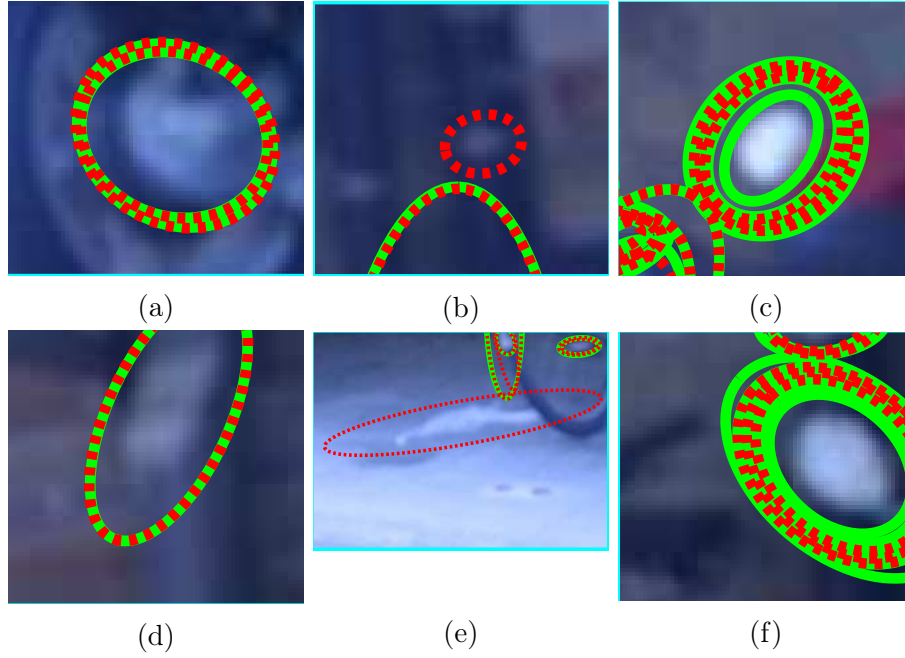


Figura 5.5: Algunas regiones detectadas en la imagen número 4 del conjunto Bikes.
(a) y (d) Regiones detectadas igual en I y ρ . (b) y (e) Regiones detectadas solo en I .
(c) y (f) Regiones detectadas más veces en ρ que en I .

RTP que en el HSI. Este efecto se debe a que el desenfoque genera cambios más suaves en el espacio RTP que en el HSI y por ende varios mínimos locales en la variación de áreas.

5.3 Desempeño de descriptores con color

De acuerdo con los resultados registrados en la sección 5.2, en la etapa de detección con MSER no se encontraron diferencias significativas al usar la componente de intensidad del espacio HSI (I) o RTP (ρ). Por esta razón se usaron las regiones detectadas con ambas formas de intensidad como puntos característicos y se encontraron los descriptores SURF y SIFT con los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP para encontrar los indicadores de desempeño R : Razón de verdaderas correspondencias, P : Precisión y $F_{0.5}$.

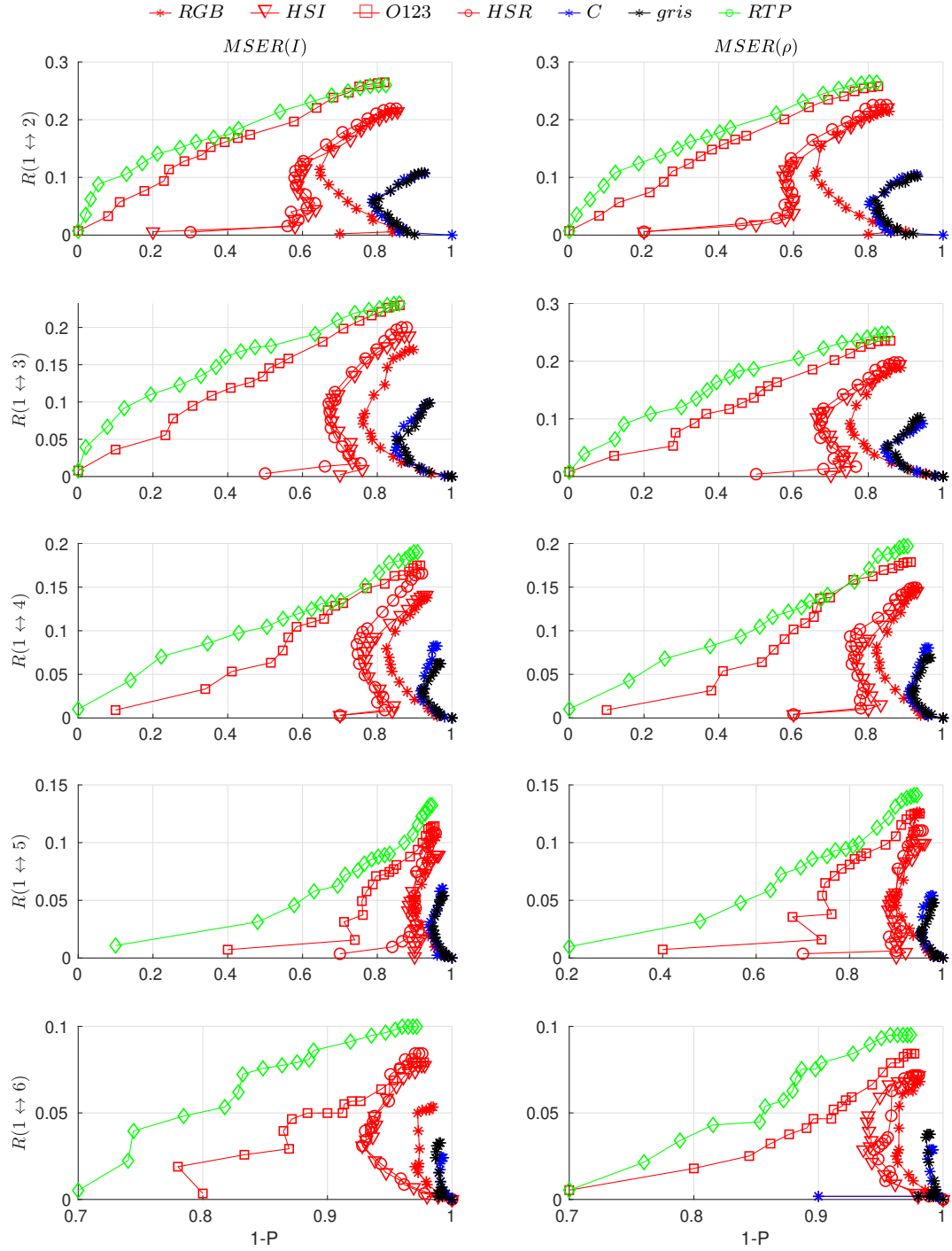


Figura 5.6: Razón de verdaderas correspondencias vs 1-Precisión para cada par del conjunto *Wall*

La Figura 5.6 muestra las curvas R vs $1 - P$ obtenidas en la etapa de descripción con

el descriptor SURF con los espacios de color RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP usando los pares estereoscópicos $1 \leftrightarrow X$ del conjunto de conjunto *Wall* ($X \in \{2, 3, 4, 5, 6\}$). Las columnas izquierda y derecha muestran las curvas obtenidas al usar los puntos de interés encontrados con el detector MSER con las componentes de intensidad del espacio de color HSI ($MSER(I)$) y RTP ($MSER(\rho)$) respectivamente. Se puede observar que con todas las representaciones evaluadas, el desempeño disminuye con la deformación de las imágenes ocasionada por el cambio de punto de vista, no obstante, el desempeño obtenido con el espacio de color RTP tiene menor disminución que con las otras representaciones. Todas las curvas R vs $1 - P$ generadas para cada par ($1 \leftrightarrow X$) formado entre la imagen de referencia 1 y la imagen X de cada conjunto de imágenes evaluado usando los descriptores SURF y SIFT se registran en los apéndices B y D respectivamente.

Se realizó una comparación cuantitativa con base en la función $F_{0.5}$ descrita en la sección 3.3.2. La Figura 5.7 muestra las curvas promedio de los valores $F_{0.5}$ ($\overline{F_{0.5}}$) asociados a cada par estereoscópico $1 \leftrightarrow X$ ($X \in 2, 3, 4, 5, 6$) del conjunto *Wall* usando el descriptor SURF con los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP con los puntos característicos detectados con la componente de intensidad en los espacios de color HSI ($MSER(I)$) y RTP ($MSER(\rho)$). En esta Figura se observa que el desempeño del descriptor SURF es mayor cuando se utiliza el espacio de color RTP en todas los pares formados en el conjunto *Wall*.

Todas curvas promedio de los valores $F_{0.5}$ para cada par ($1 \leftrightarrow X$) de cada conjunto de imágenes evaluado con los descriptores SURF y SIFT se registran en los apéndices C y E.

La Figura 5.8 muestra los valores promedio de $F_{0.5}$ en cada conjunto de imágenes usando el descriptor SURF con los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP operando con los puntos detectados por MSER con las componentes I y ρ .

Al usar SURF (Ver Figura 5.8) con el espacio de color RTP se presentan mejores

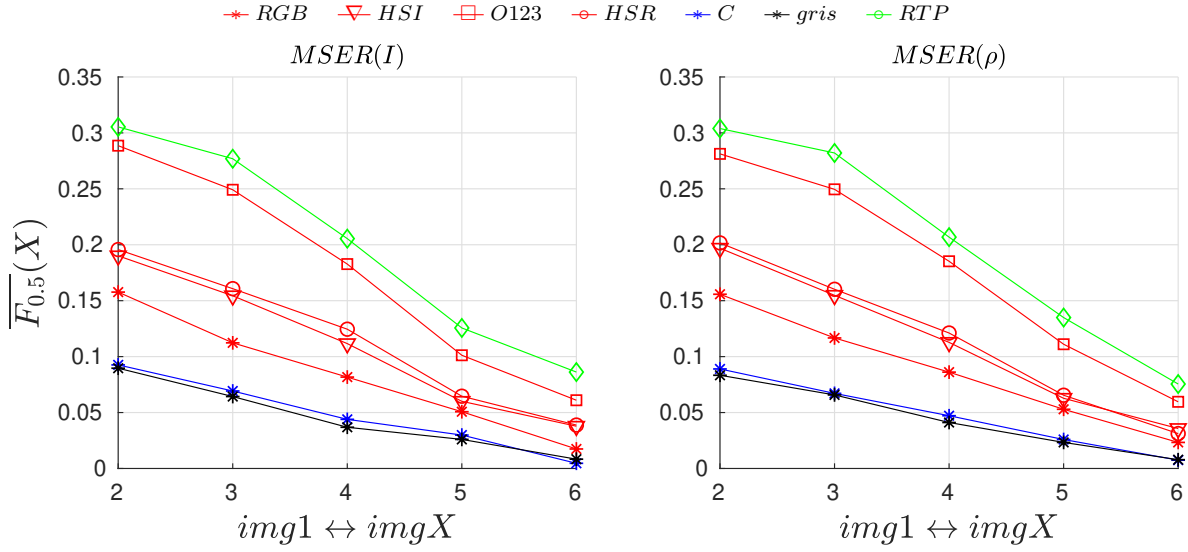


Figura 5.7: Promedio de $F_{0.5}$ con los descriptores SURF con color para cada par del conjunto Wall.

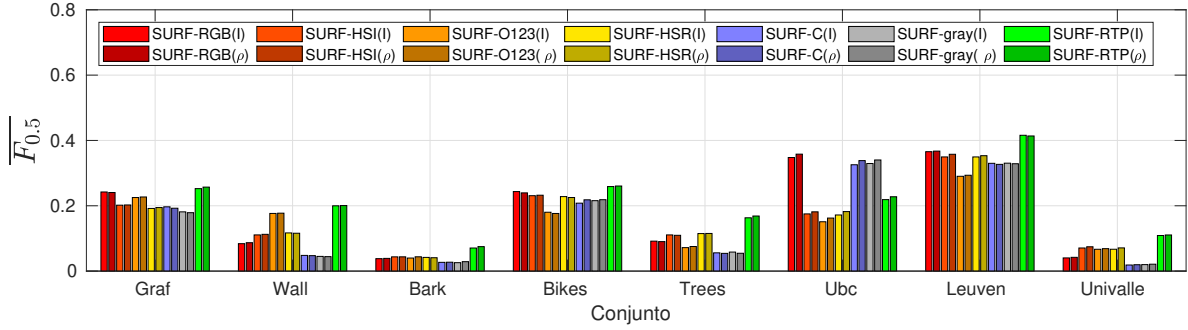


Figura 5.8: Comparación del desempeño de descriptores con color mediante los valores $F_{0.5}$ promedio por conjunto evaluado para el descriptor SURF.

resultados promedio en todos los conjuntos de imágenes evaluados, excepto en el conjunto UBC, que corresponde a las transformaciones ocasionadas por el nivel de compresión de las imágenes. Además, el desempeño es mayor al detectar los puntos característicos sobre la intensidad del espacio de RTP que la del espacio HSI.

La Figura 5.9 muestra los valores promedio de $F_{0.5}$ en cada conjunto de imágenes usando el descriptor SIFT con los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP operando con los puntos detectados por MSER con las componentes I y ρ .

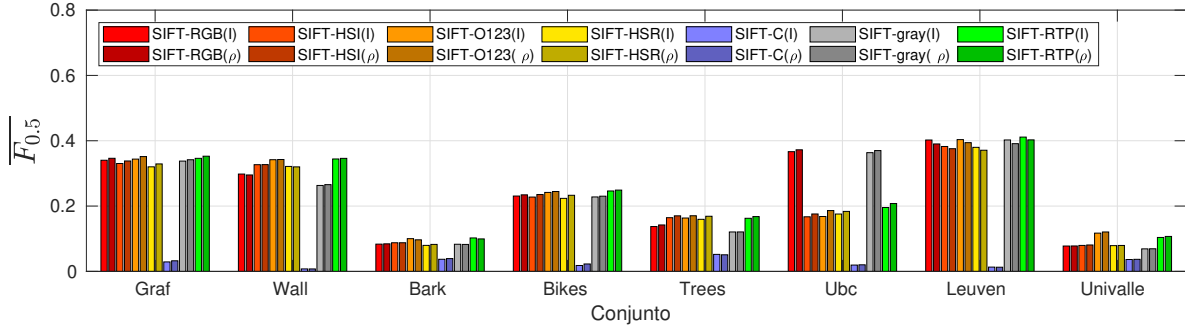


Figura 5.9: Comparación del desempeño de descriptores con color mediante los valores $F_{0.5}$ promedio por conjunto evaluado para el descriptor SIFT.

Al usar SIFT (Ver Figura 5.9) con el espacio de color RTP se presentan mejores resultados promedio en los conjuntos *Graf*, *Wall*, *Bark*, *Bikes* y *Leuven*. En el conjunto *Trees* se destacan los espacios HSI, de colores oponentes $O_1O_2O_3$ y el espacio esférico HSR. En el conjunto *UBC* se destaca el espacio RGB y los tonos de gris. Finalmente, en el conjunto *Univalle* se destaca el espacio de colores oponentes $O_1O_2O_3$. Para el descriptor SIFT, el desempeño es mayor al detectar los puntos característicos sobre la intensidad del espacio de RTP que la del espacio HSI.

5.3.1 Aporte de RTP a la descripción con SURF

Para realizar un análisis más detallado del aporte del espacio de color RTP al descriptor SURF en cada tipo de transformación evaluada, se realiza una prueba ANOVA (Análisis de varianza unifactorial) con los valores $F_{0.5}$ asociados a cada combinación detector–descriptor teniendo como hipótesis nula que los valores medios de $F_{0.5}$ son iguales para todas las combinaciones evaluadas. Esta prueba se realizó sobre cada tipo de transformación de manera individual, con todos los conjuntos sin incluir el efecto de compresión para tener en cuenta las condiciones típicas en un vuelo fotogramétrico y sobre todos los conjuntos de imágenes para tener en cuenta todos los efectos combinados.

El Cuadro 5.2 lista los valores p de cada ANOVA que indican, con una certeza del 95 %, que en cada transformación existe al menos una combinación detección–descripción que presenta diferencias estadísticamente significativas ($P < 0.05$) respecto a otra, y los valores medios de $F_{0.5}$ obtenidos para cada

combinación evaluada. Se han resaltado los mejores desempeños para cada tipo de transformación y forma de intensidad utilizada en la detección con MSER. Puede observarse que el espacio de color RTP permite obtener mejores desempeños con todas las transformaciones evaluadas excepto en el caso de la pérdida de información debido a compresión, donde el espacio de color de mejor desempeño es el RGB.

MSER	SURF	RotEsc	Desenfoco	PuntoVista	Iluminacion	Aereas	Compre.
I	RGB	3.80	16.76	16.31	36.56	4.03	34.78
	HSI	4.36	17.08	15.63	34.95	7.08	17.51
	O123	4.00	12.60	20.09	29.03	6.68	15.11
	HSR	4.19	17.15	15.43	34.95	6.69	17.18
	C	2.68	13.19	12.24	33.00	1.86	32.56
	Gris	2.58	13.69	11.32	33.04	1.99	32.94
	RTP	7.07	21.10	22.62	41.58	10.89	21.89
ρ	RGB	3.89	16.49	16.38	36.74	4.22	35.81
	HSI	4.37	17.09	15.75	35.77	7.45	18.22
	O123	4.38	12.59	20.21	29.35	6.87	16.24
	HSR	4.09	17.03	15.52	35.33	7.10	18.22
	C	2.72	13.63	11.99	32.68	1.95	33.83
	Gris	2.88	13.66	11.15	32.87	2.11	34.02
	RTP	7.49	21.46	22.89	41.37	11.06	22.76
valor - p		<0.05	<0.05	<0.05	<0.05	<0.05	<0.05

Cuadro 5.2: Resultados estadísticos del análisis ANOVA para el descriptor SURF.

Se han resaltado los valores medios más altos para cada cambio analizado.

En la Figura 5.10 se ilustran, para cada transformación evaluada, los intervalos de confianza, con una certeza del 95 %, de los valores medios de $F_{0.5}$ asociados a los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP, en combinación con el descriptor SURF y la forma de intensidad sobre la que se realiza la detección de puntos característicos con MSER.

Una inspección visual de los intervalos de confianza (Ver Figura 5.10) de la media de la métrica $F_{0.5}$ permite afirmar que:

- Ante cambios debidos a rotación y escala, el espacio RTP brinda un mejor desempeño que los demás con diferencias estadísticamente significativas (Ver Figura 5.10a).
- Ante cambios debidos a enfoque, el espacio RTP presenta mejoras con diferencias estadísticamente significativas respecto al espacio de colores oponentes $O_1O_2O_3$,

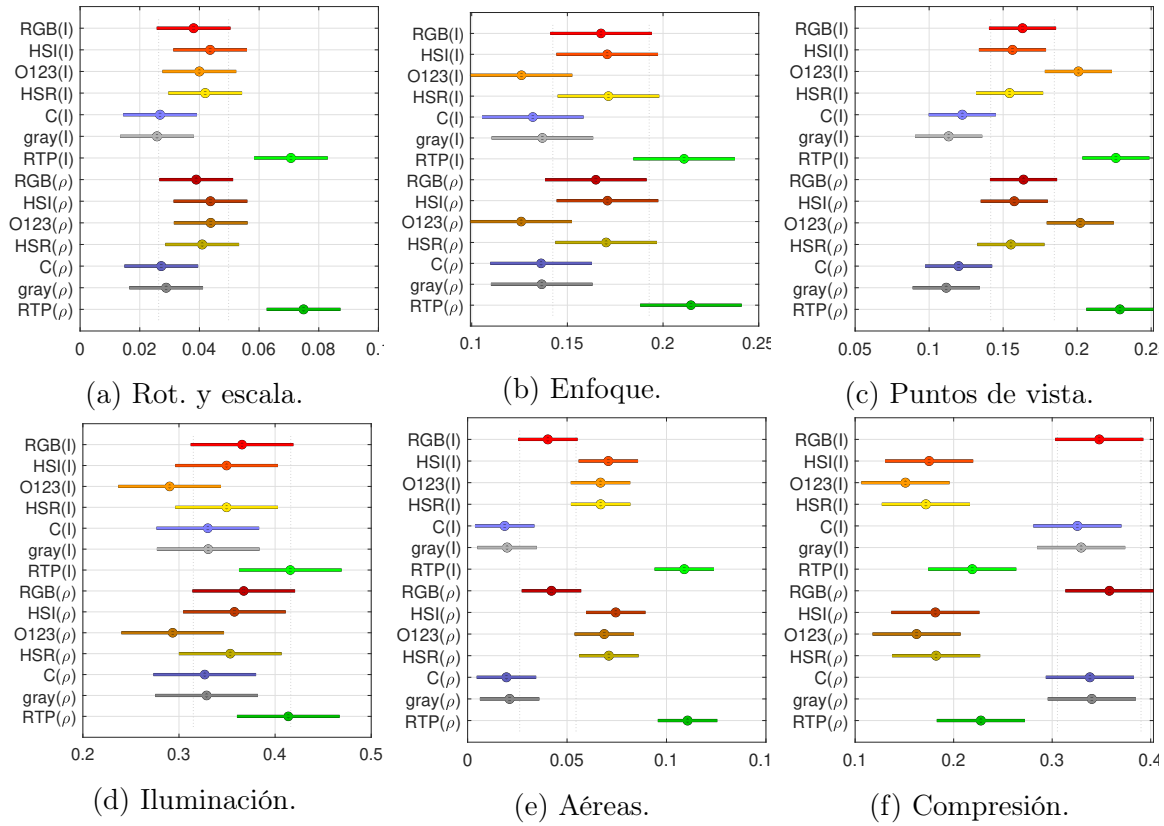


Figura 5.10: Intervalos de confianza de valores medios de $F_{0.5}$ para cada transformación con descriptores SURF.

las tonalidades de gris y la función de invariantes cromáticas C . Además, RTP muestra leves mejoras respecto a los espacios RGB, HSI y HSR (Ver Figura 5.10b).

- Ante cambios de puntos de vista, el espacio RTP brinda un mejor desempeño, con diferencias estadísticamente significativas, respecto a todos los espacios y representaciones evaluadas, excepto con el espacio de colores oponentes $O_1O_2O_3$, contra el cual no presenta diferencias estadísticamente significativas pero sí leves mejoras (Ver Figura 5.10c).
- Ante cambios de iluminación, el espacio de color RTP brinda un mejor desempeño que los demás, pero solo presenta diferencias estadísticamente significativas respecto al espacio de colores oponentes $O_1O_2O_3$ (Ver Figura 5.10d).
- Para la correspondencia entre las imágenes aéreas el espacio de color RTP proporciona un mejor desempeño, con diferencias estadísticamente significativas,

que los espacios de color $O_1O_2O_3$, RGB, HSR, HSI, las tonalidades de gris y la función de invariantes cromáticas C (Ver Figura 5.10e).

- Ante cambios debidos a compresión, el espacio de color RTP brinda leves mejoras respecto a los espacios HSI, HSR y de colores oponentes. Su desempeño es menor, con diferencias estadísticamente significativas que el obtenido con el espacio RGB, los tonos de gris y la función invariante cromática C (Ver Figura 5.10f).
- En todos los casos, los desempeños mejoran levemente si se detectan los puntos característicos sobre la intensidad del espacio RTP.

El Cuadro 5.3 lista los valores p del análisis ANOVA sobre todas las transformaciones típicas en vuelos fotogramétricos e incluyendo el efecto de la compresión. Esta muestra con una certeza del 95 % que existe al menos una combinación detección—descripción que presenta diferencias estadísticamente significativas ($P < 0.05$) respecto a las demás. El Cuadro también lista los valores medios de $F_{0.5}$ obtenidos para cada combinación evaluada. Se han resaltado los mejores desempeños para cada consideración y forma de intensidad utilizada en la detección con MSER. Puede observarse que, en términos generales, el espacio de color RTP permite obtener mejores desempeños en el proceso de correspondencia basada en el descriptor SURF.

La Figura 5.11a ilustra los intervalos de confianza de valores medios de $F_{0.5}$ que se obtienen al considerar todas las transformaciones típicas en los vuelos fotogramétricos. La Figura 5.11b ilustra los intervalos de confianza de valores medios de $F_{0.5}$ que se obtienen al considerar también la pérdida de información relacionada a la compresión. Al tener en cuenta los efectos de todas las transformaciones, incluyendo la compresión, el espacio de color RTP brinda, con diferencias estadísticamente significativas, un desempeño mejor que los espacios HSI, HSR, de colores oponentes $O_1O_2O_3$, la representación en tonos de gris y con la función invariante cromática C . Las mejoras son leves respecto al espacio RGB (Ver Figura 5.11b). Sin embargo, si se tienen en cuenta solamente las deformaciones de imagen típicas en un vuelo fotogramétrico, el espacio de color RTP brinda un mejor desempeño que los demás espacios de color y representaciones cromáticas con diferencias estadísticamente significativas (Ver Figura 5.11a).

MSER	SURF	General	Gen. + Compr.
<i>I</i>	RGB	15.79	18.16
	HSI	15.97	16.17
	O123	15.01	15.02
	HSR	15.86	16.02
	C	12.63	15.12
	Gris	12.52	15.07
	RTP	21.00	21.11
<i>ρ</i>	RGB	15.80	18.30
	HSI	16.18	16.43
	O123	15.17	15.30
	HSR	15.95	16.23
	C	12.66	15.30
	Gris	12.50	15.19
	RTP	21.23	21.42
valor-p		<0.05	<0.05

Cuadro 5.3: Resultados estadísticos del análisis ANOVA general para el descriptor SURF.

5.3.2 Aporte de RTP a la descripción con SIFT

Para realizar un análisis más detallado del aporte del espacio de color RTP al descriptor SIFT en cada tipo de transformación evaluada, se realiza una prueba ANOVA (Análisis de varianza unifactorial) con los valores $F_{0.5}$ asociados a cada combinación detector–descriptor teniendo como hipótesis nula que los valores medios de $F_{0.5}$ son iguales para todas las combinaciones evaluadas. Esta prueba se realizó sobre cada tipo de transformación de manera individual, con todos los conjuntos sin incluir el efecto de compresión para tener en cuenta las condiciones típicas en un vuelo fotogramétrico y sobre todos los conjuntos de imágenes para tener en cuenta todos los efectos combinados.

El Cuadro 5.4 lista los valores p de cada ANOVA que indican, con una certeza del 95 %, que en cada transformación existe al menos una combinación detección–descripción que presenta diferencias estadísticamente significativas ($P < 0.05$) respecto a otra, y los valores medios de $F_{0.5}$ obtenidos para cada combinación evaluada. Se han resaltado los mejores desempeños para cada tipo de transformación y forma de intensidad utilizada en la detección con MSER. Puede

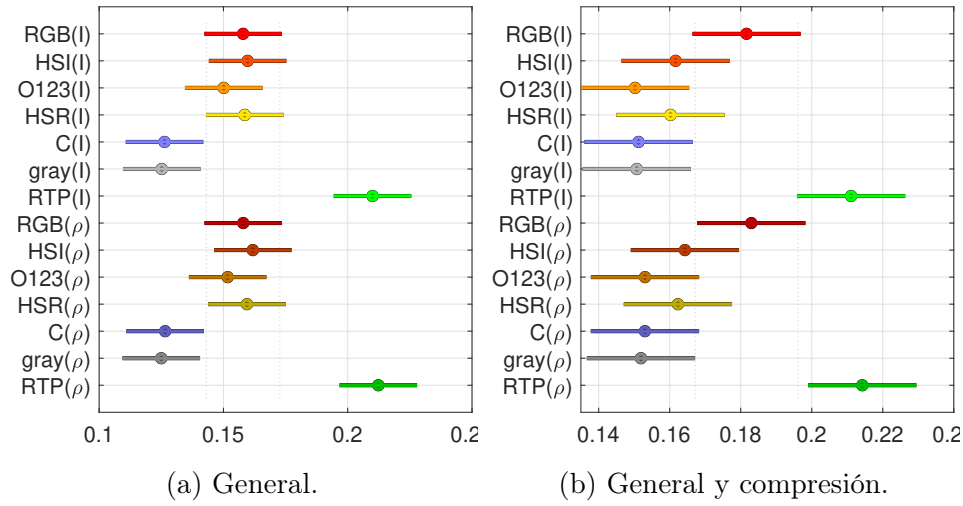


Figura 5.11: Intervalos de confianza de valores medios de $F_{0.5}$ para todas las transformaciones evaluadas con el descriptor SURF.

observarse que el espacio de color RTP permite obtener mejores desempeños con todas las transformaciones evaluadas excepto en el caso de la pérdida de información debido a compresión y distorsiones en imágenes aéreas, donde los espacios de mejor desempeño son RGB y de colores oponentes $O_1O_2O_3$ respectivamente.

En la Figura 5.12 se ilustran, para cada transformación evaluada, los intervalos de confianza, con una certeza del 95 %, de los valores medios de $F_{0.5}$ asociados a los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, la función de invariante cromática C , la capa gris de intensidad y el espacio propuesto RTP, en combinación con el descriptor SIFT y la forma de intensidad sobre la que se realiza la detección de puntos característicos con MSER.

Una inspección visual de los intervalos de confianza (Ver Figura 5.12) de la media de la métrica $F_{0.5}$ permite afirmar que:

- Ante cambios debidos a rotación y escala, enfoque, puntos de vista e iluminación, el espacio RTP brinda un mejor desempeño que los demás, sin embargo, solo presenta diferencias estadísticamente significativas con la representación mediante la función cromática invariante C (Ver Figura 5.12a).
- Para la correspondencia entre las imágenes aéreas, el espacio de color RTP

MSER	SIFT	RotEsc	Desenfoque	PuntoVista	Iluminacion	Aereas	Compre.
I	RGB	8.35	18.40	31.92	40.23	7.77	36.66
	HSI	8.77	19.60	32.85	38.24	7.95	16.72
	O123	10.02	20.26	34.29	40.35	11.72	16.81
	HSR	7.96	19.15	32.08	37.98	7.89	17.56
	C	3.71	3.51	1.83	1.31	3.64	1.94
	Gris	8.31	17.43	30.04	40.24	6.90	36.35
	RTP	10.24	20.44	34.52	41.12	10.39	19.56
ρ	RGB	8.44	18.83	32.08	39.02	7.79	37.23
	HSI	8.77	20.27	33.25	37.57	8.09	17.57
	O123	9.67	20.75	34.70	39.42	12.07	18.62
	HSR	8.26	20.09	32.45	37.10	7.94	18.37
	C	3.92	3.68	2.00	1.28	3.70	2.02
	Gris	8.26	17.55	30.37	39.11	6.93	36.99
	RTP	9.95	20.84	34.94	40.27	10.71	20.78
valor - p		<0.05	<0.05	<0.05	<0.05	<0.05	<0.05

Cuadro 5.4: Resultados estadísticos del análisis ANOVA para el descriptor SIFT. Se han resaltado los valores medios más altos para cada cambio analizado.

proporciona un mejor desempeño que el resto de representaciones exceptuando el espacio de colores oponentes. Además, presenta diferencias estadísticamente significativas solamente con la representación mediante la función cromática invariante C (Ver Figura 5.12e).

- Ante cambios debidos a compresión, el espacio de color RTP muestra leves mejoras respecto a los espacios HSI, HSR y de colores oponentes. También muestra mejoras con diferencias estadísticamente significativas respecto a la representación cromática C . Sin embargo, su desempeño es menor, con diferencias estadísticamente significativas, que el obtenido con el espacio RGB y los tonos de gris (Ver Figura 5.12f).
- En todos los casos, los desempeños mejoran levemente si se detectan los puntos característicos sobre la intensidad del espacio RTP.

El Cuadro 5.5 lista los valores p del análisis ANOVA sobre todas las transformaciones típicas en vuelos fotogramétricos e incluyendo el efecto de la compresión. Esta muestra, con una certeza del 95 %, que existe al menos una combinación detección–descripción que presenta diferencias estadísticamente significativas ($P < 0.05$) respecto a las demás. El Cuadro también lista los valores medios de $F_{0.5}$ obtenidos para cada combinación evaluada. Se han resaltado los mejores desempeños para cada consideración y forma

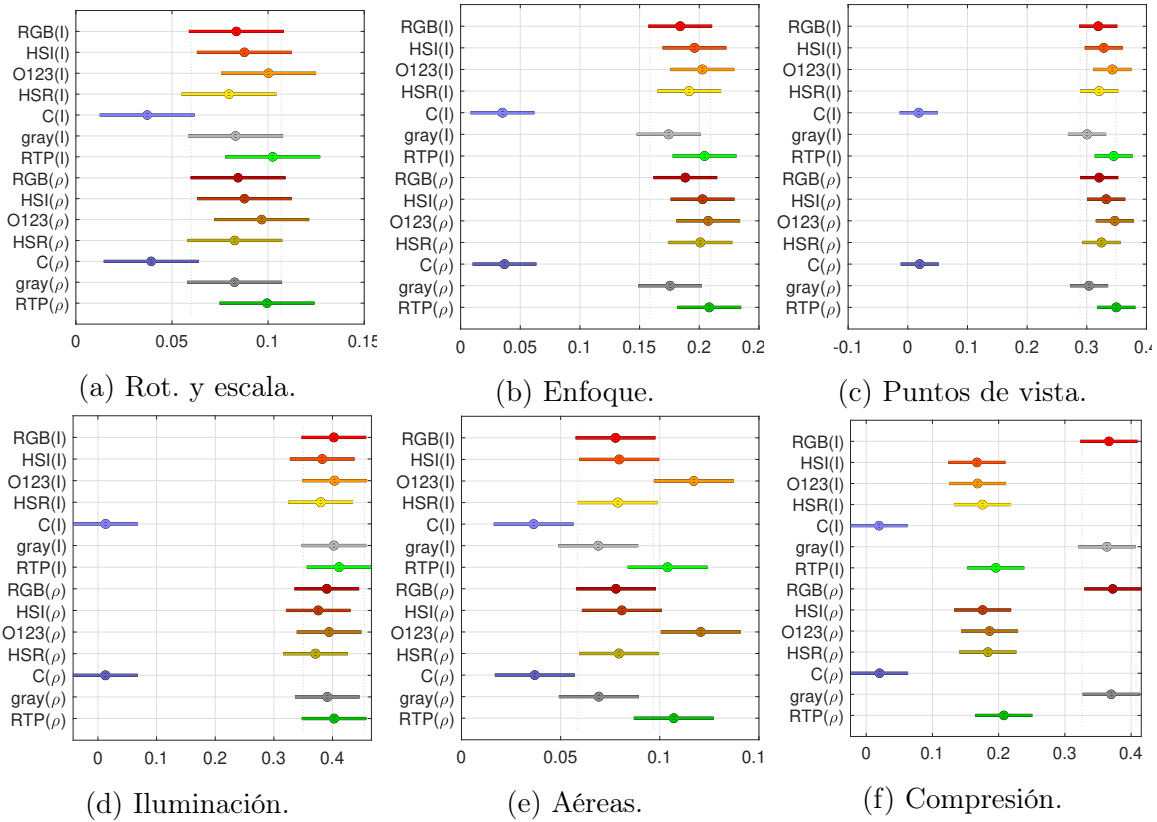


Figura 5.12: Intervalos de confianza de valores medios de $F_{0.5}$ para cada transformación con descriptores SIFT.

de intensidad utilizada en la detección con MSER. Puede observarse que, en términos generales, el espacio de color RTP permite obtener mejores desempeños en el proceso de correspondencia basada en el descriptor SIFT, aunque el espacio de color RGB se destaca si se tiene en cuenta el efecto de la compresión.

La Figura 5.13a ilustra los intervalos de confianza de valores medios de $F_{0.5}$ que se obtienen al considerar todas las transformaciones típicas en los vuelos fotogramétricos. La Figura 5.13b ilustra los intervalos de confianza de valores medios de $F_{0.5}$ que se obtienen al considerar también la pérdida de información relacionada a la compresión. En las condiciones normales de un vuelo fotogramétrico, el espacio de color RTP brinda un desempeño levemente mejor que los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, HSR y la representación en tonos de gris. La mejora tiene diferencias estadísticamente significativas respecto a la función invariante cromática C (Ver Figura 5.13a). Por otra parte, si se incluye el efecto de compresión, el espacio de

MSER	SIFT	General	Gen. + Compr.
I	RGB	22.43	24.20
	HSI	22.84	22.07
	O123	24.45	23.50
	HSR	22.33	21.73
	C	2.76	2.66
	Gris	21.49	23.34
	RTP	24.52	23.90
ρ	RGB	22.44	24.29
	HSI	23.07	22.38
	O123	24.58	23.83
	HSR	22.63	22.09
	C	2.89	2.78
	Gris	21.45	23.39
	RTP	24.64	24.16
valor-p		<0.05	<0.05

Cuadro 5.5: Resultados estadísticos del análisis ANOVA general para el descriptor SIFT.

color RGB se destaca como el de mejor desempeño pero no tiene diferencias estadísticamente significativas respecto al espacio propuesto RTP (Ver Figura 5.13b).

5.4 Construcción de un mosaico con imágenes aéreas

Con un subconjunto de las imágenes a color capturadas en un vuelo fotogramétrico realizado el 17 de enero de 2019 a una altura de 81 m en una zona plana de la Universidad del Valle, se construyó un mosaico utilizando la metodología descrita en el capítulo 4. Como la metodología propuesta permite encontrar correspondencias entre dos imágenes, se utilizó iterativamente para encontrar las correspondencias entre imágenes consecutivas del conjunto de imágenes tomando la primera como referencia para el sistema de coordenadas del mosaico y ajustando una transformación afín acumulativa para reprojectar cada imagen al sistema de referencia. La Figura 5.14 muestra resaltada en azul la zona de estudio y en verde la trayectoria diseñada, cuyo punto final se señala con el indicador rojo.

Las correspondencias identificadas entre un par de imágenes pueden ser falsas. Por esta razón se impone una restricción que obliga a que las correspondencias se ubiquen

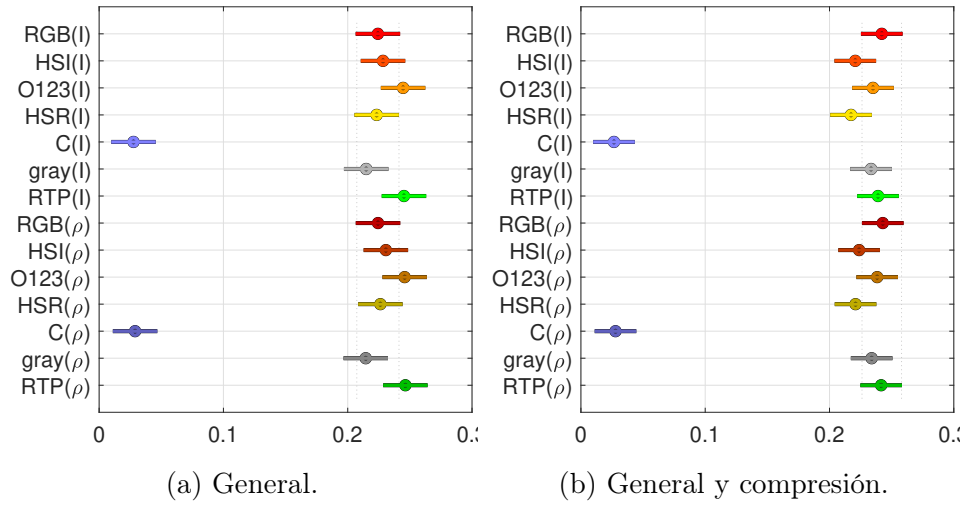


Figura 5.13: Intervalos de confianza de valores medios de $F_{0.5}$ para todas las transformaciones evaluadas con el descriptor SIFT.

sobre las líneas epipolares estimadas con muestras de los puntos identificados como correspondientes. La restricción epipolar permite eliminar falsas correspondencias y calcular la relación geométrica entre las imágenes evaluadas.

Las Figuras 5.15 y 5.16 muestran los pares de puntos que fueron identificados como correspondientes en un dos imágenes del vuelo mencionado usando los descriptores SIFT y SURF, respectivamente, con las imágenes en tonos de gris y los espacios HSI, RGB y RTP.

Las Figuras 5.15a, 5.15c, 5.15e y 5.15g ilustran las correspondencias estimadas con el descriptor SIFT mientras las Figuras 5.15b, 5.15d, 5.15f y 5.15h ilustran las correspondencias estimadas que cumplen con la restricción epipolar y se identificaron como verdaderas correspondencias con el descriptor SIFT.

El Cuadro 5.6 muestra la cantidad de puntos que fueron identificados como correspondientes entre las imágenes apareadas, la cantidad de puntos que cumplen con la restricción epipolar y la proporción de verdaderas correspondencias en este par de imágenes. Puede observarse que el descriptor SIFT operando sobre el espacio RTP encuentra más correspondencias (16) y más verdaderas correspondencias (15) que cuando opera sobre las imágenes en tonos de gris, o los espacios HSI y RTP.

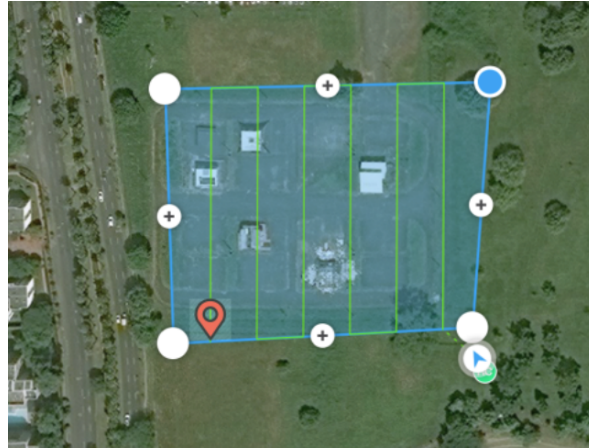


Figura 5.14: Ruta del vuelo para generar el mosaico.

Representación	Correspondencias Iniciales	Correspondencias Verdaderas	Proporción
Gris	9	8	0.89
HSI	6	6	1.00
RGB	9	8	0.89
RTP	16	15	0.94

Cuadro 5.6: Resumen de correspondencias encontradas entre dos imágenes aéreas con SIFT.

Las Figuras 5.16a, 5.16c, 5.16e y 5.16g ilustran las correspondencias estimadas con el descriptor SURF mientras las Figuras 5.16b, 5.16d, 5.16f y 5.16h ilustran las correspondencias estimadas que cumplen con la restricción epipolar y se identificaron como verdaderas correspondencias con el descriptor SURF.

El Cuadro 5.7 muestra la cantidad de puntos que fueron identificados como correspondientes entre las imágenes apareadas, la cantidad de puntos que cumplen con la restricción epipolar y la proporción de verdaderas correspondencias en este par de imágenes. Puede observarse que el descriptor SURF operando sobre el espacio RTP más verdaderas correspondencias (22) y mayor proporción entre verdaderas correspondencias y correspondencias encontradas que cuando opera sobre las imágenes en tonos de gris, o los espacios HSI y RTP.

Finalmente, la Figura 5.17 muestra un mosaico creado con la metodología propuesta usando cinco imágenes del vuelo fotogramétrico. Puede observarse que la transformación

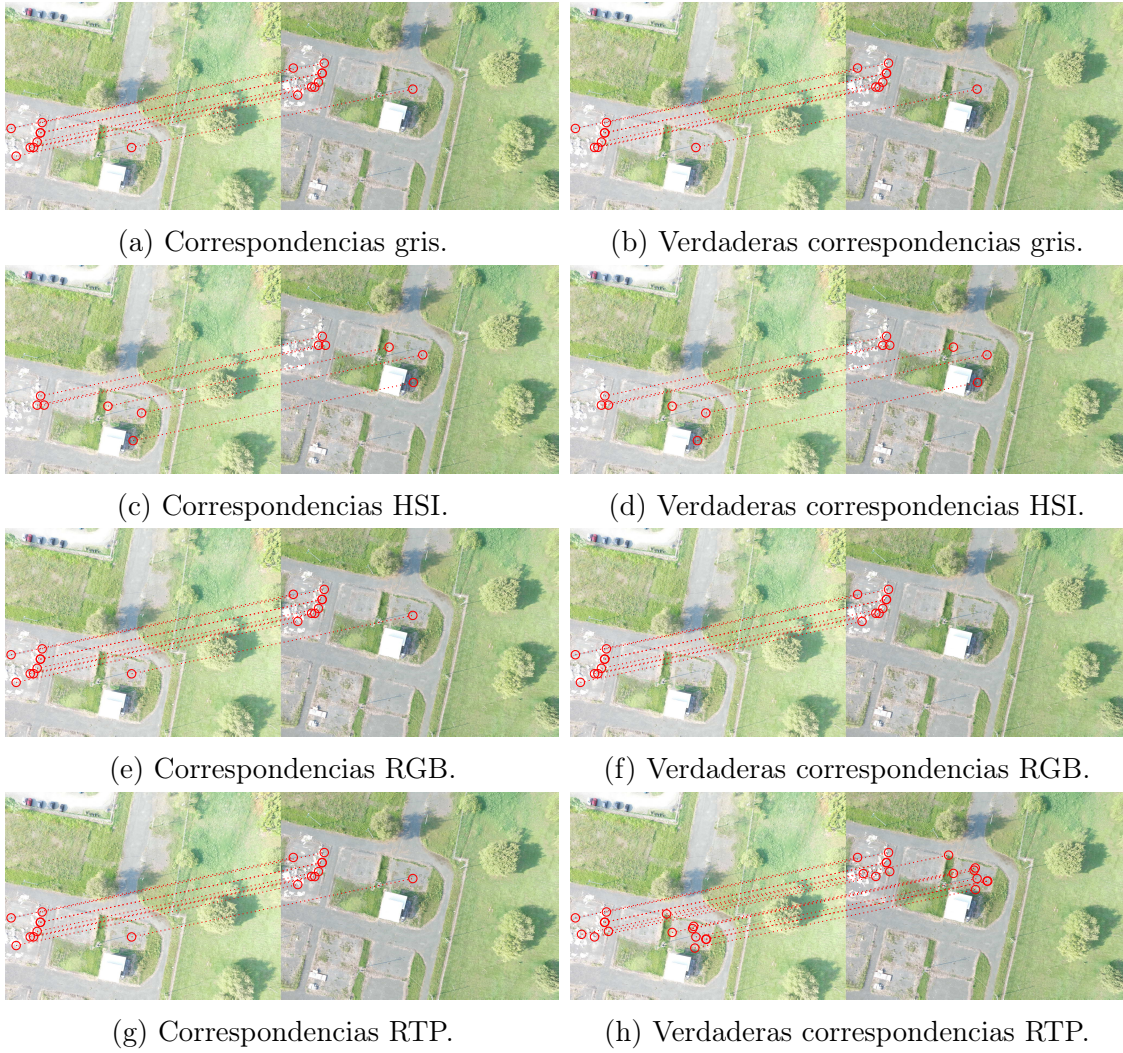


Figura 5.15: Pares y verdaderas correspondencias con el descriptor SIFT.

estimada para imágenes consecutivas genera pequeños errores al integrar las imágenes al mosaico. Estos errores se deben a que la transformación afín ajustada para una imagen se obtiene únicamente con base en la imagen inmediatamente anterior y no las imágenes con puntos de captura más cercanos en el espacio geográfico.

5.5 Discusión

A partir de los resultados obtenidos en la evaluación al detector MSER mediante el análisis de *Repetibilidad* se puede afirmar que:

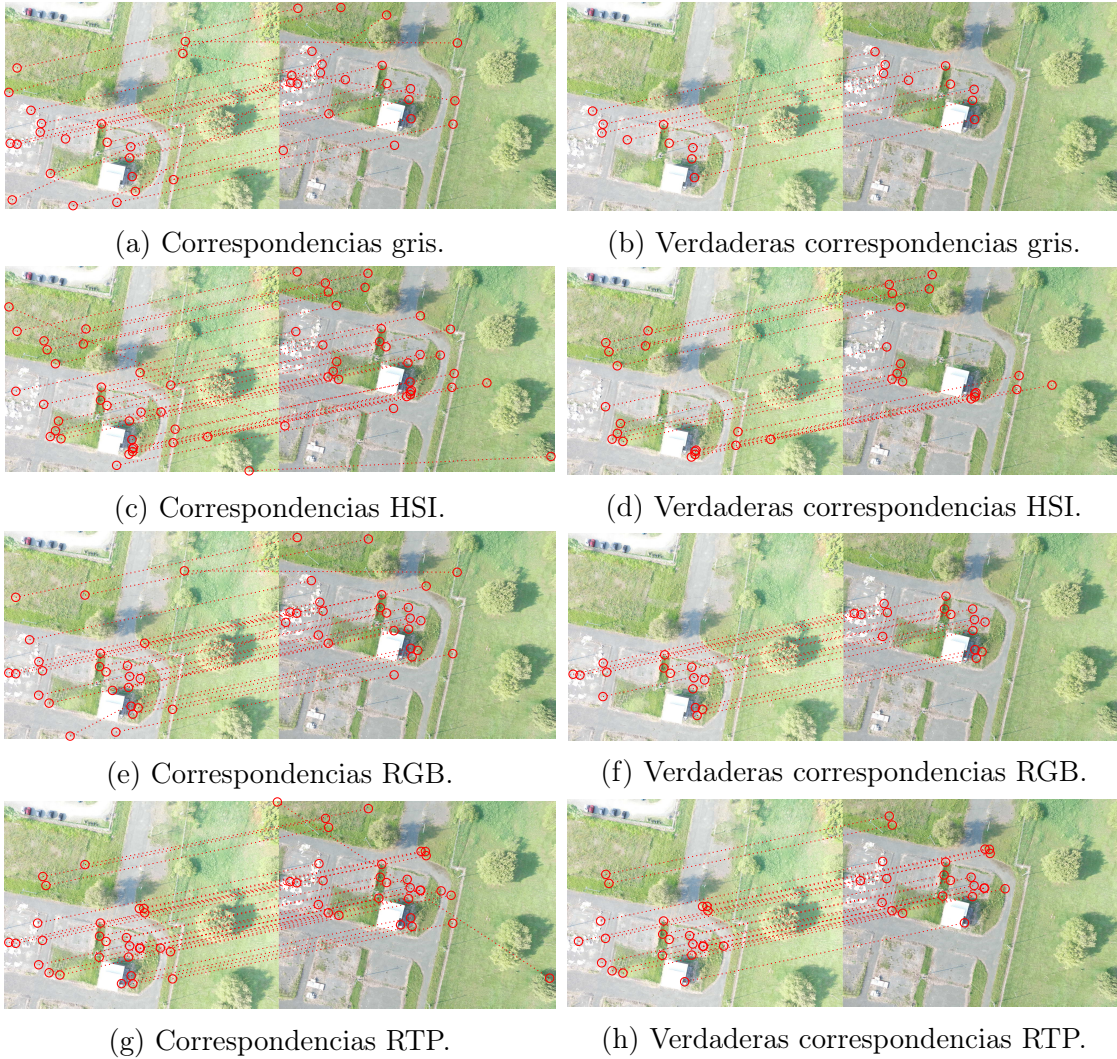


Figura 5.16: Pares y verdaderas correspondencias con el descriptor SURF.

- Al utilizar el detector MSER como la base de la etapa de detección de puntos característicos no se encontraron diferencias significativas en *Repetibilidad* al usar la información de intensidad en los espacios de color RTP (56.02 % en promedio) y HSI (56.14 % en promedio).
- La detección de puntos sobre la intensidad de RTP encuentra múltiples regiones asociadas al mismo objeto. Esto se debe a la detección de varios niveles de estabilidad de área por cambios de intensidad más suaves en zonas de poco cambio cromático, lo que crea una baja razón de cambio de área respecto a los niveles de umbralización.

Representación	Correspondencias Iniciales	Correspondencias Verdaderas	Proporción
Gris	21	9	0.43
HSI	31	17	0.55
RGB	25	15	0.60
RTP	29	22	0.76

Cuadro 5.7: Resumen de correspondencias encontradas entre dos imágenes aéreas con SURF.

A partir de los resultados obtenidos en la evaluación de descripción mediante el análisis de *Razón de verdaderas correspondencias*, *Precisión* y valores $F_{0.5}$ se puede afirmar que:

- Al detectar los puntos sobre la intensidad de RTP se obtienen leves mejoras en el desempeño de todos descriptores basados en histogramas de gradientes que fueron evaluados. Esto debido a que la multiplicidad de puntos con distintas áreas asociadas se ve penalizada por el criterio de repetibilidad utilizado en el análisis de la etapa de detección mientras favorece en la correspondencia al dar más posibilidades a un punto de encontrar su correspondiente en la otra imagen.
- La representación mediante función invariante cromática C no es apropiada para la creación de los descriptores SIFT. Esta representación introdujo los desempeños más bajos respecto a las demás representaciones con diferencias estadísticamente significativas en la evaluación del descriptor SIFT.
- En términos generales, los descriptores SIFT y SURF con el espacio de color RTP presentan mejor desempeño promedio respecto a los descriptores obtenidos con los espacios de color RGB, HSI, de colores oponentes $O_1O_2O_3$, HSI, las imágenes en tonos de gris y representadas con la función invariante a color C .



Figura 5.17: Mosaico construido con la primera linea de vuelo usando SURF con RTP.

6. Conclusiones

- Se desarrolló un procedimiento para la puesta en correspondencia entre imágenes mediante métodos semiglobales que permite demostrar las implicaciones que tienen los espacios de color en la detección y descripción de puntos característicos, y su contribución al desempeño de la puesta en correspondencia. El análisis del efecto del espacio de color en los procedimientos de puesta en correspondencia no ha sido abordado y los avances se concentran en las etapas de detección y descripción.

Este procedimiento incorpora una etapa de *conversión entre espacios de color*, una etapa de *detección* que emplea MSER operando sobre la componente de intensidad del espacio de color, una etapa de *descripción* basada en histogramas de gradientes sobre los tres componentes del espacio de color, y una etapa de *correspondencia* que emplea como criterio el vecino más cercano en el espacio de características.

El análisis de los espacios de color RGB, HSI, su equivalente en coordenadas esféricas y el espacio de colores ponentes $O_1O_2O_3$ puso en evidencia la existencia de problemas de consistencia de dichos espacios con los supuestos euclidianos sobre los que se basa el cálculo de gradientes en la etapa de descripción, por lo que se propuso un nuevo espacio de color, denominado RTP, que se sobrepone completamente a la limitante mencionada.

- El método semiglobal de puesta en correspondencia utilizado para el análisis, se priorizó sobre los métodos locales y globales, por presentar una mejor relación entre precisión y tiempo de ejecución, necesitar menos recursos computacionales que los métodos globales, y no requerir conocimiento a priori de la geometría de captura como el caso general de los métodos locales.
- La etapa de detección se implementó con el detector MSER debido a que reporta el mejor desempeño en términos de *Repetibilidad* en la detección de regiones asociadas a puntos de interés, es invariante ante cambios de intensidad, conserva la adyacencia entre las regiones cuando hay ausencia de oclusión y permite la detección en diversas escalas.

- Considerando sus altos desempeños en términos de *Razón de verdaderas correspondencias* y *Precisión*, en la etapa de descripción se emplearon los descriptores SIFT y SURF sobre las tres componentes de los espacios de color y la capa de tonos de gris y la representación invariante cromática C . Estos descriptores se basan en histogramas de gradientes y usan las diferencias entre valores de píxeles dentro de vecindades alrededor de los puntos característicos para describir cómo se distribuyen dichas diferencias respecto a la orientación y posición en cada vecindad. Así, se obtienen descriptores robustos ante algunas alteraciones fotométricas y geométricas en las imágenes.
- El espacio de color RTP se propuso con el fin de preservar el poderío de la información cromática, eliminar las inconsistencias en las métricas de distancia cromática y mejorar el desempeño de la puesta en correspondencia. RTP es un espacio de color no perceptual basado en coordenadas esféricas de tres dimensiones donde, a diferencia de los espacios RGB, HSI y colores oponentes $O_1O_2O_3$, no están presentes los problemas de consistencia con los supuestos euclidianos que inducen errores en el cálculo de diferencias y gradientes.
- Se emplearon como indicadores de desempeño las medidas de *Repetibilidad*, *Razón de verdaderas correspondencias*, *Precisión* y valores $F_{0.5}$, siguiendo un protocolo de pruebas ampliamente usado por la comunidad científica [Mikolajczyk and Schmid, 2005, Burghouts and Geusebroek, 2009, Liao and Liu, 2010, Canclini et al., 2013, Lee and Park, 2017]. Para evaluar el desempeño de la etapa de detección se emplea la *Repetibilidad* y para la descripción se emplea $F_{0.5}$, que se obtiene de las medidas de *Razón de verdaderas correspondencias* y *Precisión*. El desempeño de los métodos de puesta en correspondencia con los diferentes espacios de color y representaciones cromáticas se comparó mediante un análisis ANOVA de los valores $F_{0.5}$ con un nivel de certeza del 95 %. Este procedimiento estadístico se propone para sobreponerse a la subjetividad de la comparación mediante inspección visual de curvas R vs $1 - P$ y al posible sesgo en la elección del máximo $F_{0.5}$.
- Debido a que ANOVA es una generalización de contraste de igualdad de medias para muestras independientes, permite comparar simultáneamente el efecto de diversas metodologías sobre un indicador de desempeño con un nivel de certeza determinado. Por lo cual se propone que esta comparación estadística sea incluida

en trabajos futuros para el análisis comparativo de metodologías para la puesta en correspondencia. En este estudio, ANOVA permitió comparar el efecto de los espacios de color sobre los valores $F_{0.5}$ en un conjunto de muestras con una certeza del 95 %.

- Los resultados obtenidos permiten establecer que el espacio RTP introduce mejoras en el desempeño de las etapas de detección y descripción en la metodología semiglobal de puesta en correspondencia respecto a los espacios RGB, HSI, de colores oponentes $O_1O_2O_3$, el espacio esférico HSR, las imágenes en tonos de gris y representada por la función invariante cromática C .
- No se encontraron diferencias significativas en la *Repetibilidad* del detector MSER al usar la información de intensidad en los espacios de color RTP (56.02 % en promedio) y HSI (56.14 % en promedio), sin embargo, se identificó que al detectar los puntos sobre la intensidad de RTP se obtiene una mejora de 0.003 en el desempeño (valor medio de $F_{0.5}$) de la etapa de descripción con descriptores basados en histogramas de gradientes. Esto debido a que algunos puntos se detectan varias veces con distintas áreas asociadas. Esta multiplicidad se ve penalizada por el criterio de *Repetibilidad* utilizado en el análisis de la etapa de detección pero favorece al proceso de correspondencia al dar más posibilidades a un punto de encontrar su correspondiente en la otra imagen. La multiplicidad en detección que se obtiene al usar el espacio RTP se debe a cambios de intensidad más suaves en zonas de poco cambio cromático, lo que crea una baja razón de cambio de área respecto a los niveles de umbralización. Esto repercute en la detección de varios niveles de estabilidad de área y por ende zonas repetidas.
- Se puede afirmar con un 95 % de certeza que el espacio de color RTP mejora, con diferencias estadísticamente significativas, el desempeño del descriptor SURF con respecto a las imágenes en tonos de gris, la función invariante a color C [Abdel-Hakim and Farag, 2006], los espacios de color RGB, Colores oponentes $O_1O_2O_3$, HSI y HSR en condiciones de cambios en rotación y escala ($\overline{F_{0.5}} = 0.075$), enfoque ($\overline{F_{0.5}} = 0.215$), puntos de vista ($\overline{F_{0.5}} = 0.229$), en iluminación ($\overline{F_{0.5}} = 0.414$) y capturas aéreas ($\overline{F_{0.5}} = 0.111$).
- No se encontraron diferencias estadísticamente significativas al usar el descriptor SIFT con los espacios de color RGB, HSI y su equivalente esférico HSR, de

colores oponentes $O_1O_2O_3$ y la representación en tonos de gris, sin embargo se identificaron leves mejoras al usar el espacio de color RTP en condiciones de cambios en rotación y escala ($\overline{F_{0.5}} = 0.100$), enfoque ($\overline{F_{0.5}} = 0.208$), puntos de vista ($\overline{F_{0.5}} = 0.349$), en iluminación ($\overline{F_{0.5}} = 0.403$).

- Aunque en la correspondencia de imágenes aéreas con el descriptor SIFT, el espacio de colores oponentes $O_1O_2O_3$ mostró mejor desempeño ($\overline{F_{0.5}} = 0.121$) que RTP ($\overline{F_{0.5}} = 0.107$), no se encontraron diferencias estadísticamente significativas entre estos.
- Los descriptores contruidos con el espacio de color RTP presentan mejor desempeño promedio respecto a los otros espacios de color evaluados para todas las transformaciones consideradas en este estudio, excepto la transformación de pérdida de información debida compresión de imágenes, ante la cual el espacio de color de mejores prestaciones es el RGB. En el caso de la compresión, cualquier transformación de espacio de color ejecutada resulta en una propagación de errores desde el espacio RGB original.

7. Perspectivas de trabajos futuros

Tomando como base los aportes desarrollados para el problema de puesta en correspondencia, se plantean las siguientes perspectivas para desarrollos futuros:

- La incorporación de otros espacios de color al análisis comparativo de desempeños en el problema de puesta en correspondencia.
- Inclusión de descriptores binarios al estudio comparativo.
- Nueva metodología para encontrar histogramas de gradientes cromáticos calculados en el espacio de color RTP con las leyes de la trigonometría esférica.
- Adaptación del espacio RTP para encontrar su homólogo espacio de color perceptual y comparar su desempeño ante distintas aplicaciones respecto a los espacios de color de *CIE*.
- Descripción global, usando el espacio de color RTP, como etapa previa a la correspondencia entre puntos característicos para eliminar la búsqueda de correspondencias entre imágenes poco relacionadas.
- Aplicación de reconstrucción 3D y generación de ortomosaicos basada en los descriptores enriquecidos con el espacio RTP que tenga en cuenta los metadatos de posición que tienen las imágenes y los use, junto a las correspondencias encontradas, para optimizar las posiciones de captura en el espacio de la escena.
- Determinación de la combinación apropiada de componentes de distintos espacios de color para mejorar el desempeño en correspondencia.

8. Productos

Durante el desarrollo de esta investigación se obtuvieron los siguientes productos:

- **Classical and neural models for binocular stereoscopic reconstruction (2015):** Artículo presentado en STSIVA 2015 donde se evidenciaba el problema de correspondencia como central en la reconstrucción 3D.
- **Visual Terrain Identification and Surface Inclination Estimation for Improving Human Locomotion with a Lower-Limb Prosthetic (2018):** Artículo presentado en EMBC 2018 donde se empiezan a explotar los descriptores locales con color en el problema de detectar coberturas.
- **Un espacio esférico para el color:** Presentación en el Simposio de investigaciones 2018 con la primera aproximación del espacio RTP.
- **Técnicas y aplicaciones del procesamiento digital de imágenes en los ámbitos de la ingeniería civil (2018):** Presentación en el III ciclo Nuevas tendencia del conocimiento en la Ingeniería Civil. En este se presentaron los avances y posibles aplicaciones en la ingeniería civil.
- **Espacio No Perceptual de Color Esférico RTP: Aplicación en Discriminación de Imágenes Generadas por Computador.** Artículo sometido.

Referencias

- [Abdel-Hakim and Farag, 2006] Abdel-Hakim, A. E. and Farag, A. A. (2006). CSIFT: A SIFT descriptor with color invariant characteristics. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:1978–1983.
- [Abraham et al., 2013] Abraham, E., Mishra, S., Tripathi, N., and Sukumaran, G. (2013). HOG descriptor based registration (A new image registration technique). In *2013 15th International Conference on Advanced Computing Technologies, ICACT 2013*.
- [Ackermann et al., 2014] Ackermann, J., Langguth, F., Fuhrmann, S., Kuijper, A., and Goesele, M. (2014). Multi-View Photometric Stereo by Example. In *3DV*, pages 259–266.
- [Alarcon and Dalmau, 2014] Alarcon, T. and Dalmau, O. (2014). Color Categorization Models for Color Image Segmentation. In *Advances in Low-Level Color Image Processing*, pages 303–327. Springer.
- [Alegre and Pajares, 2016] Alegre, E. and Pajares, G. (2016). *Conceptos y Métodos en Visión por Computador*, volume 1.
- [Alonso Pérez, 2009] Alonso Pérez, M. A. (2009). *Espacios de Color RGB, HSI y sus Generalizaciones a n-Dimensiones por*. PhD thesis, Instituto Nacional de Astrofísica, Óptica y Electrónica.
- [Barandiaran et al., 2013] Barandiaran, I., Cortes, C., Nieto, M., Graña, M., and Ruiz, O. E. (2013). A New Evaluation Framework and Image Dataset for Keypoint Extraction and Feature Descriptor Matching. In *8th International Conference on Computer Vision Theory and Applications (VISAPP 2013)*, number 3, pages 1–6.
- [Bay et al., 2008] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359.

- [Berrocso et al., 2003] Berrocso, M., Ramírez, M. E., Enríquez-Salamanca, J. M., and Pérez Peña, A. (2003). *Notas y apuntes de trigonometria esferica y astronomia de posicion*. Universidad de Cadiz, Servicio de Publicaciones, Cadiz.
- [Bethmann and Luhmann, 2014] Bethmann, F. and Luhmann, T. (2014). Object-based Multi-Image Semi-Global Matching - Concept and first results. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5(June):93–100.
- [Bradley et al., 2008] Bradley, D., Boubekeur, T., and Heidrich, W. (2008). Accurate multi-view reconstruction using robust binocular stereo and surface meshing. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*.
- [Brown et al., 2005] Brown, M., Szeliski, R., and Winder, S. (2005). Multi-Image Matching using Multi-Scale Oriented Patches. Technical Report December 2004.
- [Burghouts and Geusebroek, 2009] Burghouts, G. J. and Geusebroek, J. M. (2009). Performance evaluation of local colour invariants. *Computer Vision and Image Understanding*, 113(1):48–62.
- [Cai et al., 2005] Cai, X., Zhou, D., Li, G., and Zhuang, Z. (2005). A stereo matching algorithm based on color segments. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3372–3377.
- [Calakli et al., 2012] Calakli, F., Ulusoy, A. O., Restrepo, M. I., Taubin, G., and Mundy, J. L. (2012). High Resolution Surface Reconstruction from Multi-view Aerial Imagery. In *3D Imaging, Modeling, Processing, Visualization and Transmission, 3DIMPVT*, pages 25–32.
- [Calderon et al., 2013] Calderon, F., Parra, C., and Nino, C. (2013). Evaluation of Mean, Gaussian and S&G aggregation windows in stereo correspondence under presence of noise. In *Symposium of Signals, Images and Artificial Vision - STSIVA*, pages 1–5. IEEE.
- [Calonder et al., 2010] Calonder, M., Lepetit, V., Strecha, C., and Fua, P. (2010). BRIEF: Binary Robust Independent Elementary Features. In Daniilidis, K., Maragos, P., and Paragios, N., editors, *European Conference on Computer Vision*, volume 6314 of *Lecture Notes in Computer Science*, pages 778–792, Berlin, Heidelberg. Springer Berlin Heidelberg.

- [Canclini et al., 2013] Canclini, A., Cesana, M., Redondi, A., Tagliasacchi, M., Ascenso, J., and Cilla, R. (2013). Evaluation of low-complexity visual feature detectors and descriptors. *2013 18th International Conference on Digital Signal Processing, DSP 2013*, pages 0–6.
- [Cepeda-Negrete et al., 2017] Cepeda-Negrete, J., Sanchez-Yanez, R. E., Correa-Tome, F. E., and Lizarraga-Morales, R. A. (2017). Dark Image Enhancement Using Perceptual Color Transfer. *IEEE Access*, 6:14935–14945.
- [Chen et al., 2013a] Chen, T., Deng, Z., and Ma, J. (2013a). A spherical color model is intuitively simple and perceptually harmonious. *SPIE Newsroom*, 1:10–12.
- [Chen et al., 2013b] Chen, T., Ma, J., and Deng, Z. (2013b). Attributes of color represented by a spherical model. *Journal of Electronic Imaging*, 22(4):043032.
- [Clark et al., 2008] Clark, A. J., Green, R. D., and Grant, R. N. (2008). Perspective correction for improved visual registration using natural features. In *2008 23rd International Conference Image and Vision Computing New Zealand*, pages 1–6.
- [Colomina and Molina, 2014] Colomina, I. and Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92:79–97.
- [Correal et al., 2014] Correal, R., Pajares, G., and Ruz, J. (2014). Automatic expert system for 3D terrain reconstruction based on stereo vision and histogram matching. *Expert Systems with Applications*, 41(4):2043–2051.
- [Cruz-Mota et al., 2012] Cruz-Mota, J., Bogdanova, I. c., Paquier, B., Bierlaire, M., and Thiran, J.-P. (2012). Scale invariant feature transform on the sphere: Theory and applications. *International Journal of Computer Vision*, 98(2):217–241.
- [Cui and Ngan, 2011] Cui, C. and Ngan, K. N. (2011). Scale- and affine-invariant fan feature. *IEEE Transactions on Image Processing*, 20(6):1627–1640.
- [Do et al., 2013] Do, Y., Kim, S., Park, S., and Na, I. (2013). Variation of SIFT descriptor for affine invariant object matching. *International Journal of Software Engineering and its Applications*, 7(5):297–308.

- [Elibol et al., 2008] Elibol, A., Garcia, R., Delaunoy, O., and Gracias, N. (2008). A new global alignment method for feature based image mosaicing. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5359 LNCS(PART 2):257–266.
- [Fan et al., 2008] Fan, X., Rhody, H., and Saber, E. (2008). A novel feature enhanced MMI based registration algorithm for automated maps and images. In *International Geoscience and Remote Sensing Symposium (IGARSS)*, volume 2, pages II1108–II1111.
- [Ford and Roberts, 1998] Ford, A. and Roberts, A. (1998). Colour space conversions. *Westminster University, London*, 1998:1–31.
- [García Moreno et al., 2019] García Moreno, L. M., Díaz Paz, J. P., Loaiza Correa, H., and Restrepo, A. (2019). Thermal and visible aerial imagery (dataset).
- [Gauglitz et al., 2011] Gauglitz, S., Höllerer, T., and Turk, M. (2011). Evaluation of interest point detectors and feature descriptors for visual tracking. *International Journal of Computer Vision*, 94(3):335–360.
- [Goesele et al., 2010] Goesele, M., Ackermann, J., Fuhrmann, S., Klowy, R., Langguth, F., Mücke, P., and Ritz, M. (2010). Scene Reconstruction from Community Photo Collections. *Computer*, 43(6):48–53.
- [Gómez, 2016] Gómez, E. (2016). *Aportes A La Constancia Del Color Computacional Usando Máquinas De Aprendizaje*. PhD thesis, Universidad del Valle.
- [Haala, 2011] Haala, N. (2011). Multiray Photogrammetry and Dense Image Matching. In *Photogrammetric Week '11*, pages 185–195.
- [Hartley and Zisserman, 2004] Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition.
- [Hirschmuller, 2005] Hirschmuller, H. (2005). Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 807–814. IEEE.

- [Hirschmüller, 2008] Hirschmüller, H. (2008). Stereo processing by semiglobal matching and mutual information. *IEEE transactions on pattern analysis and machine intelligence*, 30(2):328–41.
- [Hong, 2010] Hong, W. (2010). *A Study of Fast, Robust Stereo-Matching*. PhD thesis, Massachusetts Institute of Technology.
- [Hosni et al., 2013] Hosni, A., Bleyer, M., and Gelautz, M. (2013). Secrets of adaptive support weight techniques for local stereo matching. *Computer Vision and Image Understanding*, 117(6):620–632.
- [Hsia et al., 2012] Hsia, K.-h., Lien, S.-f., and Su, J.-p. (2012). Stereo Matching Method and Height Estimation for Unmanned Helicopter. In *Machine Vision - Applications and Systems*.
- [Huang et al., 2014] Huang, D., Zhu, C., Wang, Y., and Chen, L. (2014). HSOG: a novel local image descriptor based on histograms of the second-order gradients. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 23(11):4680–95.
- [Kanade and Okutomi, 1991] Kanade, T. and Okutomi, M. (1991). A Stereo Matching Algorithm with an Adaptative Window: Theory and Experiment. In *International Conference on Robotics and Automation*, Sacramento.
- [Ke et al., 2004] Ke, Y., Sukthankar, R., and Huston, L. (2004). An Efficient Parts-based Near-duplicate and Sub-image Retrieval System. In *Proceedings of the 12th Annual ACM International Conference on Multimedia*, MULTIMEDIA '04, pages 869–876, New York, NY, USA. ACM.
- [Khan et al., 2011] Khan, N. Y., McCane, B., and Wyvill, G. (2011). SIFT and SURF performance evaluation against various image deformations on benchmark dataset. *2011 International Conference on Digital Image Computing: Techniques and Applications, DICTA 2011*, pages 501–506.
- [Lang et al., 2010] Lang, H., Lei, L., and Wang, Y. (2010). Enhanced point descriptors. In *ICALIP 2010 - 2010 International Conference on Audio, Language and Image Processing, Proceedings*, pages 677–681.

- [Latanov et al., 1991] Latanov, A., Polianskiĭ, V., and Sokolov, E. (1991). The 4-dimensional spherical color space in the monkey. *Zhurnal vysshei nervnoi deiatelnosti imeni I P Pavlova*, 41(4).
- [Lee and Park, 2017] Lee, M. H. and Park, I. K. (2017). Performance evaluation of local descriptors for maximally stable extremal regions. *Journal of Visual Communication and Image Representation*, 47:62–72.
- [Leonov et al., 2008] Leonov, Y. P., Sokolov, E. N., and Bowden, D. M. (2008). The representation of colors in spherical space. *Color Research and Application*, 33(2):113–124.
- [Levene, 1993] Levene, S. B. (1993). Color analysis based upon transformation to spherical coordinates.
- [Li and Zhang, 2012] Li, D. and Zhang, Y. (2012). A novel approach for the registration of weak affine images. *Pattern Recognition Letters*, 33(12):1647–1655.
- [Li et al., 2015] Li, S., Zhang, Y., and Wang, P. (2015). New stereo matching method based on improved bp algorithm. *INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS*, 8(1):464–479.
- [Liao and Liu, 2010] Liao, K. and Liu, G. (2010). An improved local feature descriptor based on SIFT. In *Proceedings of the 2nd International Conference on Internet Multimedia Computing and Service, ICIMCS’10*, pages 131–134.
- [Liu et al., 2010] Liu, C., Yang, J., and Feng, D. (2010). PPD: A robust low-computation local descriptor for mobile image retrieval. *KSII Transactions on Internet and Information Systems*, 4(3):305–323.
- [Liu et al., 2011] Liu, C., Yang, J., and Huang, H. (2011). P-SURF: A robust local image descriptor. *Journal of Information Science and Engineering*, 27(6):2001–2015.
- [Liu and Li, 2013] Liu, L. and Li, J.-X. (2013). Integrating color into the local features based on the stable color invariant regions for image retrieval. *Optik*, 124(17):2577–2582.
- [Lobaton et al., 2011] Lobaton, E., Vasudevan, R., Alterovitz, R., and Bajcsy, R. (2011). Robust topological features for deformation invariant image matching.

Proceedings of the IEEE International Conference on Computer Vision, pages 2516–2523.

[Lowe, 1999] Lowe, D. (1999). Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pages 1150–1157. IEEE.

[Lowe, 2004] Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110.

[Lynch et al., 2013] Lynch, S. E., Drew, M. S., and Finlayson, G. D. (2013). Colour constancy from both sides of the shadow edge. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 899–906.

[Matas et al., 2004] Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10 SPEC. ISS.):761–767.

[Mikolajczyk and Schmid, 2003] Mikolajczyk, K. and Schmid, C. (2003). A performance evaluation of local descriptors. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages II/257–II/263.

[Mikolajczyk and Schmid, 2004] Mikolajczyk, K. and Schmid, C. (2004). Scale & Affine Invariant Interest Point Detectors. *International Journal of Computer Vision*, 60(1):63–86.

[Mikolajczyk and Schmid, 2005] Mikolajczyk, K. and Schmid, C. (2005). A Performance evaluation of local descriptors. *IEEE transactions on pattern analysis and machine intelligence*, 27(10):1615–30.

[Mikolajczyk et al., 2005] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., and Gool, L. V. (2005). A Comparison of Affine Region Detectors. *International Journal of Computer Vision*, 65(1-2):43–72.

[Moreno et al., 2009] Moreno, P., Bernardino, A., and Santos-Victor, J. (2009). Improving the SIFT descriptor with smooth derivative filters. *Pattern Recognition Letters*, 30(1):18–26.

- [Oxford University, 2004] Oxford University (2004). Affine Covariant Regions Datasets.
- [Ramírez Rivera et al., 2014] Ramírez Rivera, A., Rojas Castillo, J., and Chae, O. (2014). Local Directional Texture Pattern image descriptor. *Pattern Recognition Letters*, 51:94–100.
- [Rassem and Khoo, 2011] Rassem, T. H. and Khoo, B. E. (2011). Object class recognition using combination of color SIFT descriptors. *2011 IEEE International Conference on Imaging Systems and Techniques, IST 2011 - Proceedings*, 101:290–295.
- [Shamoi et al., 2014] Shamoi, P., Inoue, A., and Kawanaka, H. (2014). Perceptual color space: Motivations, methodology, applications. *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems, SCIS 2014 and 15th International Symposium on Advanced Intelligent Systems, ISIS 2014*, pages 1354–1359.
- [Sokolov et al., 1979] Sokolov, E. N., Izmailov, C. A., Izmailova, T. V., and Zimachev, M. M. (1979). A Spherical Model For Color Vision. *Soviet Psychology*, 17(3):85–96.
- [Su et al., 2013] Su, D., Wu, J., Cui, Z., Sheng, V., and Gong, S. (2013). CGCI-SIFT: A more efficient and compact representation of local descriptor. *Measurement Science Review*, 13(3):132–141.
- [Susstrunk et al., 1999] Susstrunk, S., Buckley, R., and Swen, S. (1999). Standard RGB Color Spaces. In *Color Image Processing: Color Science, Systems and Applications*, pages 127–134.
- [Szeliski, 2001] Szeliski, R. (2001). Handling occlusions in dense multi-view stereo. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–103–I–110. IEEE Comput. Soc.
- [Szeliski, 2011] Szeliski, R. (2011). *Computer Vision: Algorithms and Applications*. Texts in Computer Science. Springer London, London.
- [Tola et al., 2008] Tola, E., Lepetit, V., and Fua, P. (2008). A fast local descriptor for dense matching. In *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*.

- [Tola et al., 2010] Tola, E., Lepetit, V., and Fua, P. (2010). DAISY: an efficient dense descriptor applied to wide-baseline stereo. *IEEE transactions on pattern analysis and machine intelligence*, 32(5):815–30.
- [Tola et al., 2011] Tola, E., Strecha, C., and Fua, P. (2011). Efficient large-scale multi-view stereo for ultra high-resolution image sets. *Machine Vision and Applications*, 23(5):903–920.
- [Trzcinski et al., 2015] Trzcinski, T., Christoudias, M., and Lepetit, V. (2015). Learning Image Descriptors with Boosting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3):597–610.
- [Tsay and Lee, 2012] Tsay, J.-r. and Lee, M.-s. (2012). Sift for dense point cloud matching and aero triangulation. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXIX, pages 69–74.
- [van de Sande et al., 2010] van de Sande, K. E. A., Gevers, T., and Snoek, C. G. M. (2010). Evaluating Color Descriptors for Object and Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596.
- [Verma and Liu, 2017a] Verma, A. and Liu, C. (2017a). *Recent Advances in Intelligent Image Search and Video Retrieval*, volume 121 of *Intelligent Systems Reference Library*. Springer International Publishing, Cham.
- [Verma and Liu, 2017b] Verma, A. and Liu, C. (2017b). SIFT Features in Multiple Color Spaces for Improved Image Classification. In *Recent Advances in Intelligent Image Search and Video Retrieval*, volume 121, pages 145–166.
- [Xiao et al., 2013] Xiao, X., Guo, B., Pan, F., and Shi, Y. (2013). Stereo Matching with Weighted Feature Constraints for Aerial Images. In *2013 Seventh International Conference on Image and Graphics*, pages 562–567. Ieee.
- [Yang et al., 2018] Yang, Y., Duan, F., and Ma, L. (2018). A rotationally invariant descriptor based on mixed intensity feature histograms. *Pattern Recognition*, 76:162–174.
- [Yoon and Kweon, 2006] Yoon, K.-J. and Kweon, I. S. (2006). Adaptive support-weight approach for correspondence search. *IEEE transactions on pattern analysis and machine intelligence*, 28(4):650–6.

- [Zhang et al., 2007] Zhang, D., Wang, W., Gao, W. b., and Jiang, S. (2007). An effective local invariant descriptor combining luminance and color information. In *Proceedings of the 2007 IEEE International Conference on Multimedia and Expo, ICME 2007*, pages 1507–1510.
- [Zhou and Payandeh, 2015] Zhou, J. and Payandeh, S. (2015). On user-defined region matching for augmented reality. In *Canadian Conference on Electrical and Computer Engineering*, volume 2015-June, pages 1414–1419.
- [Zhou and Hansen, 2016] Zhou, L. and Hansen, C. D. (2016). A Survey of Colormaps in Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(8):2051–2069.
- [Zhou et al., 2014] Zhou, W., Wang, C., Xiao, B., and Zhang, Z. (2014). SLD: A Novel Robust Descriptor for Image Matching. *IEEE Signal Processing Letters*, 21(3):339–342.

Apéndice A

Curvas de repetibilidad R en la etapa de detección basada en MSER

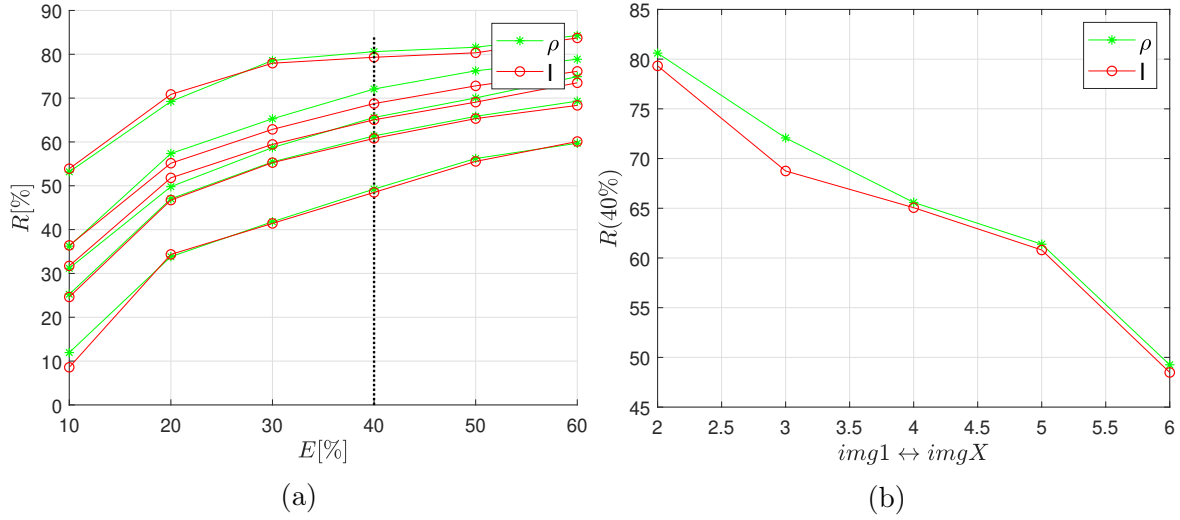


Figura A.1: Desempeño de detección ante cambios de puntos de vista en una escena estructurada. (Conjunto Graf)

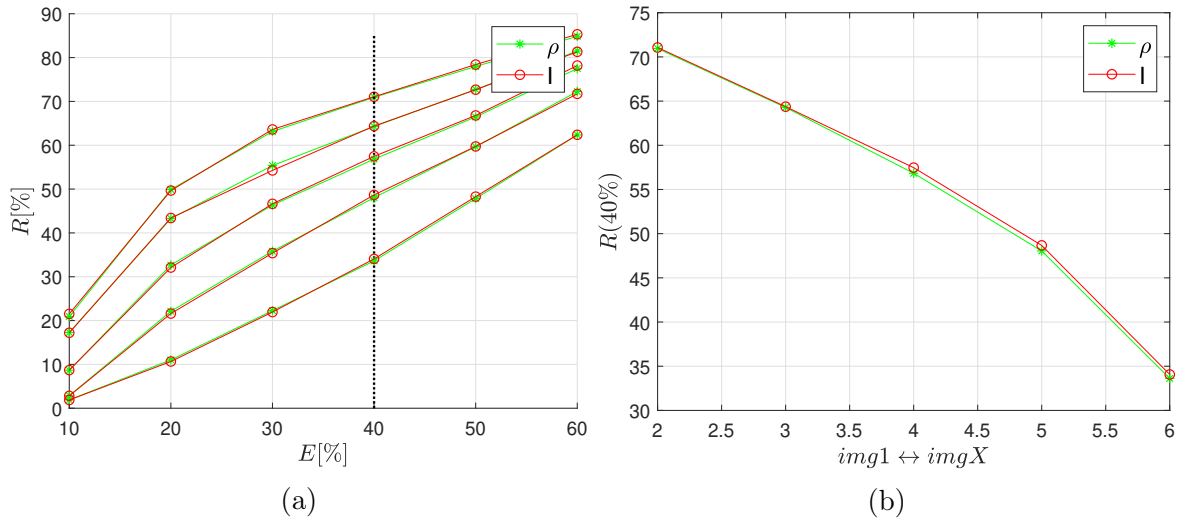


Figura A.2: Desempeño de detección ante cambios de puntos de vista en una escena con texturas. (Conjunto Wall)

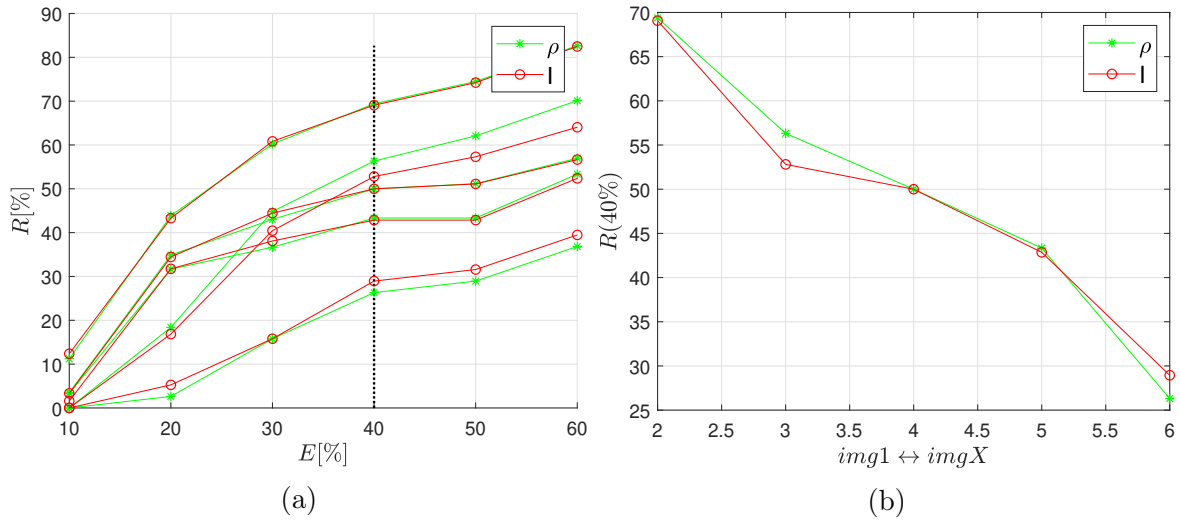


Figura A.3: Desempeño de detección ante cambios en rotación y escala en una escena con textura. (Conjunto Bark)

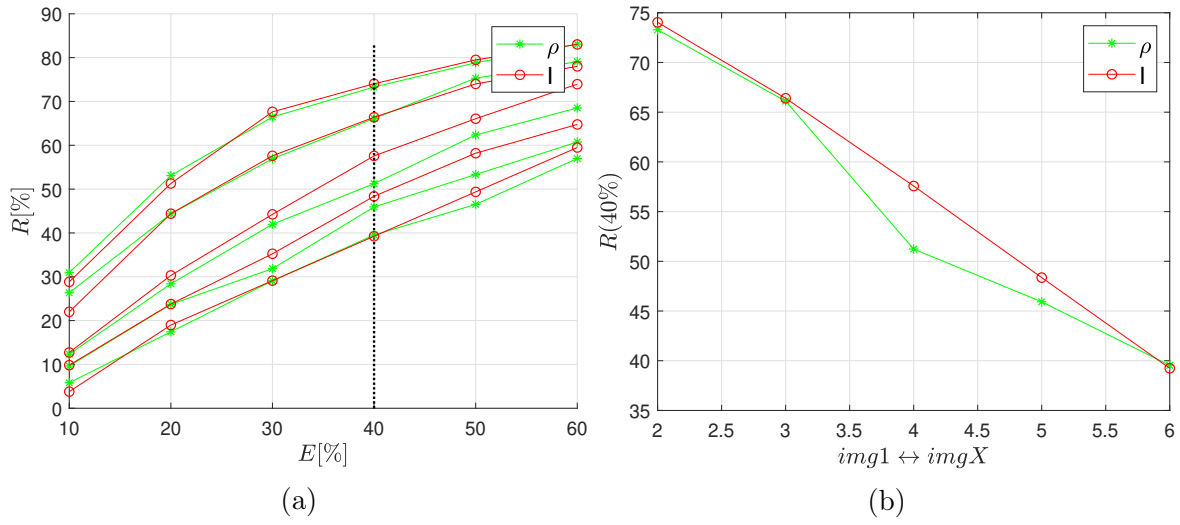


Figura A.4: Desempeño de detección ante cambios de nitidez en escena estructurada. (Conjunto Bikes)

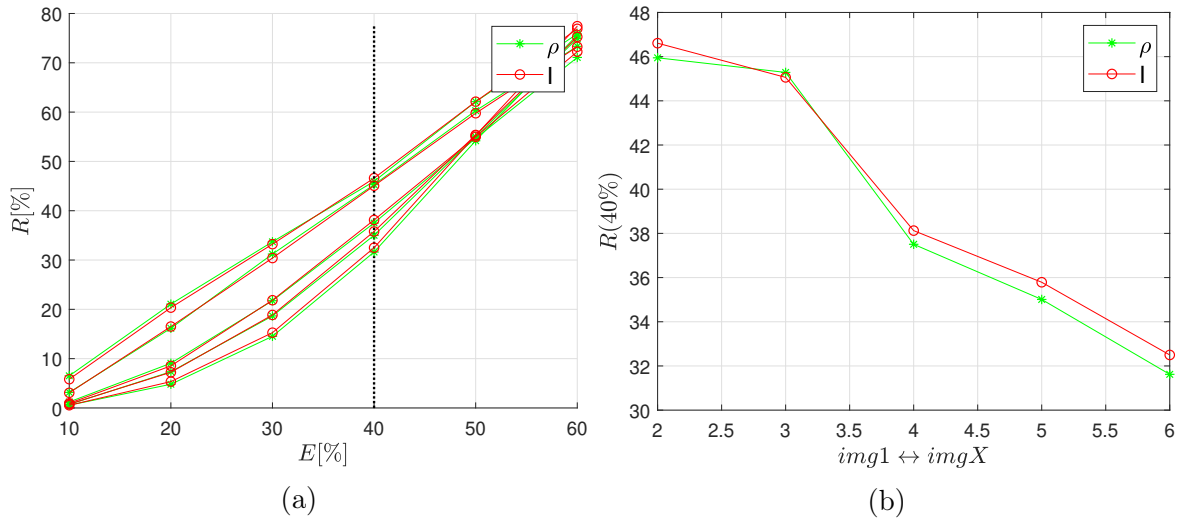


Figura A.5: Desempeño de detección ante cambios de nitidez en escena con texturas. (Conjunto Trees)

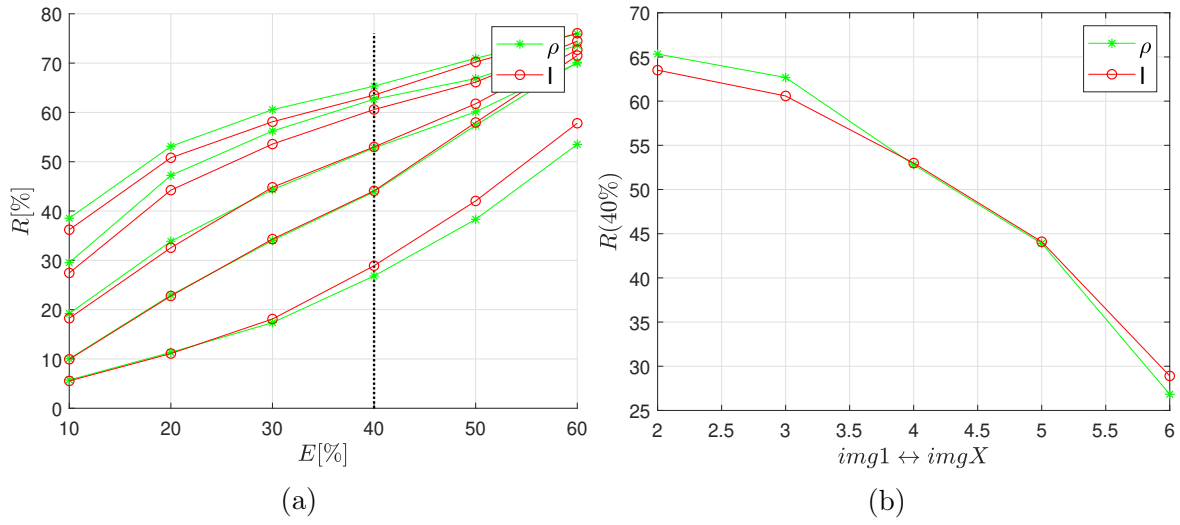


Figura A.6: Desempeño de detección ante cambios por compresión. (Conjunto UBC)
Conjunto UBC.

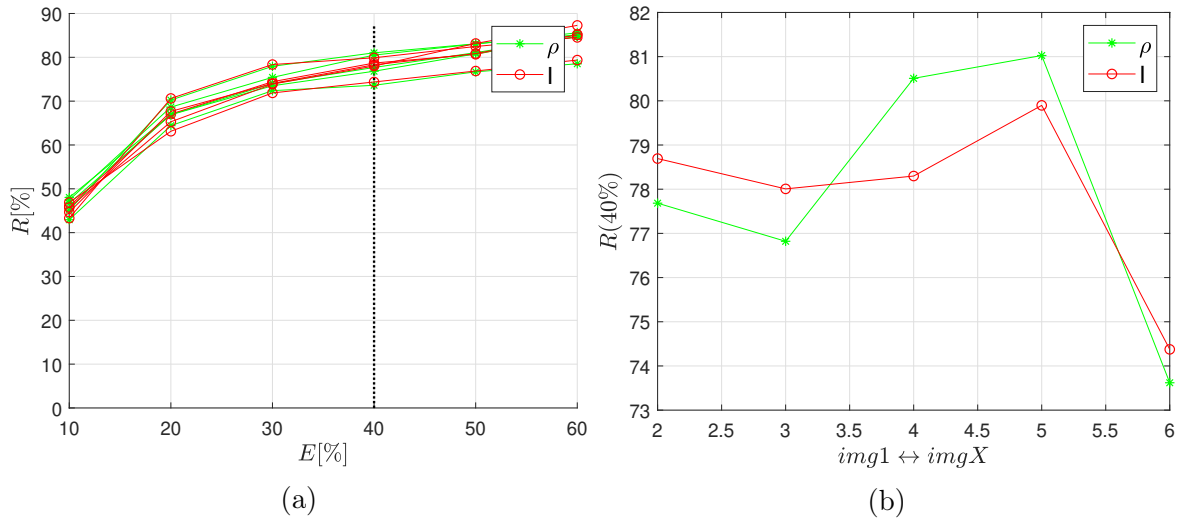


Figura A.7: Desempeño de detección ante cambios de iluminación. (Conjunto Leuven)

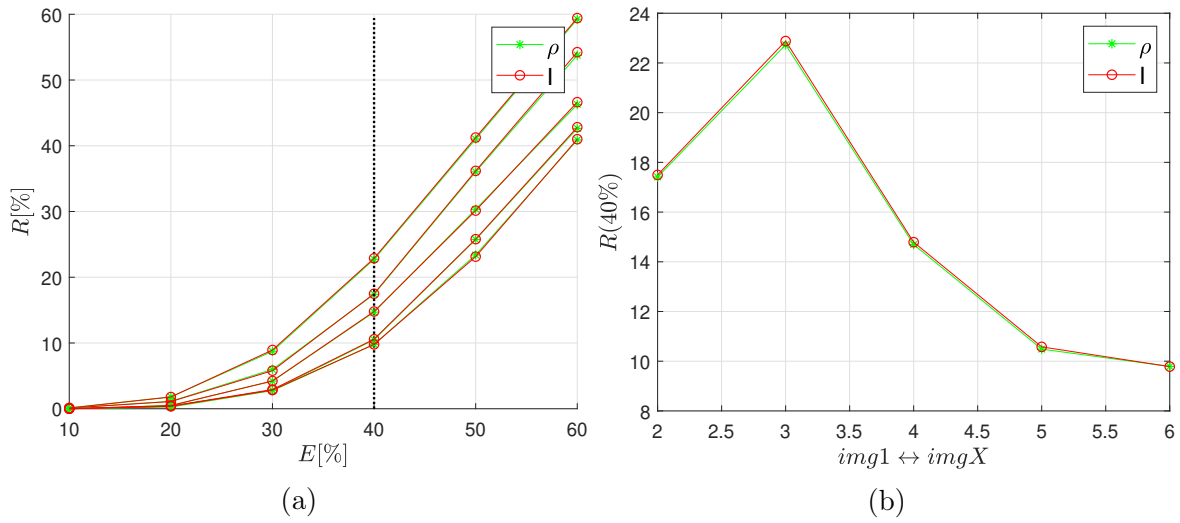


Figura A.8: Desempeño de detección con imágenes aéreas. (Conjunto Univalle)

Apéndice B

Curvas de Razón de aciertos vs 1-Precisión en la etapa de descripción basada en SURF

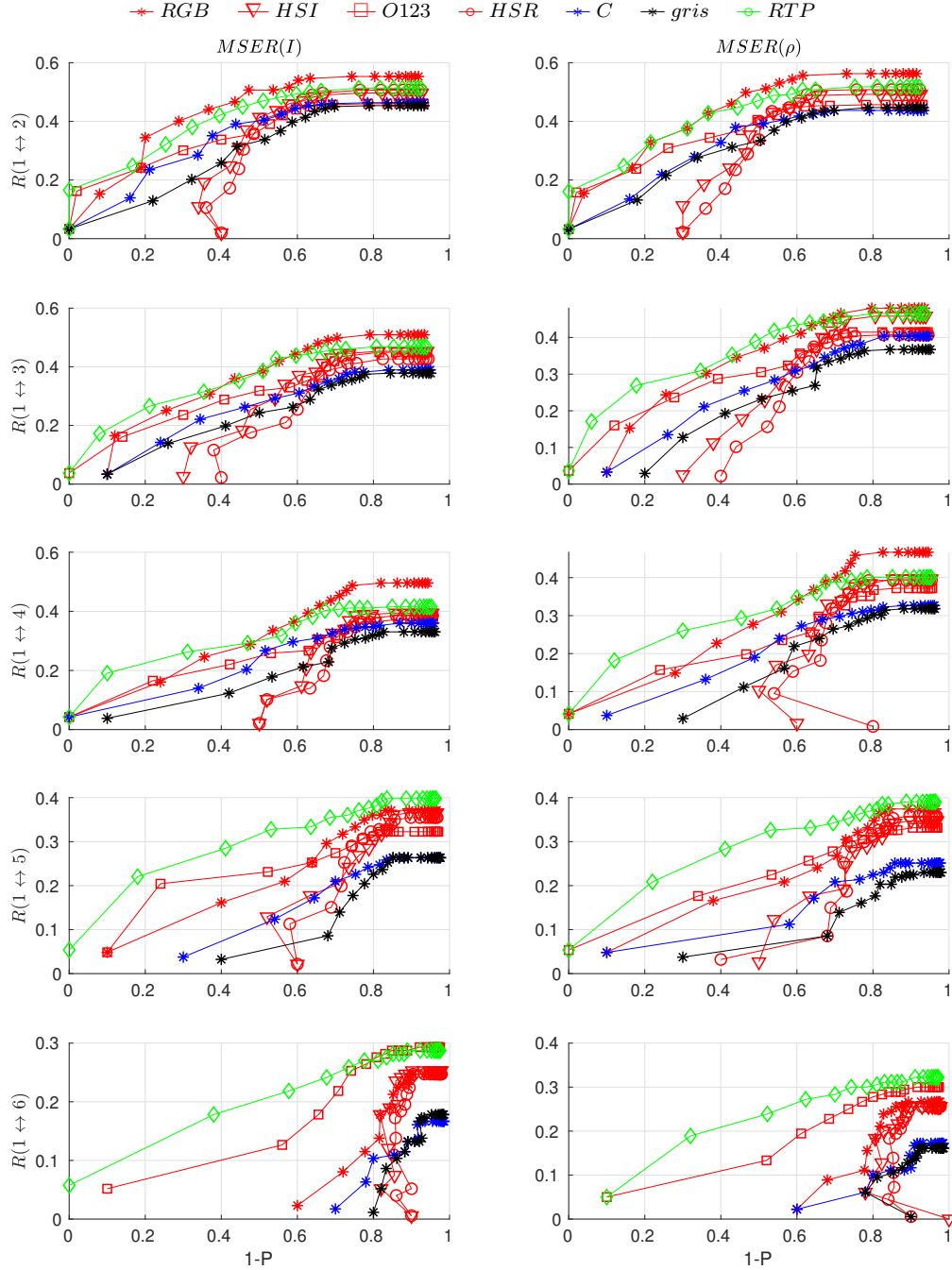


Figura B.1: R vs 1-P para el descriptor SURF en el conjunto Graffiti.

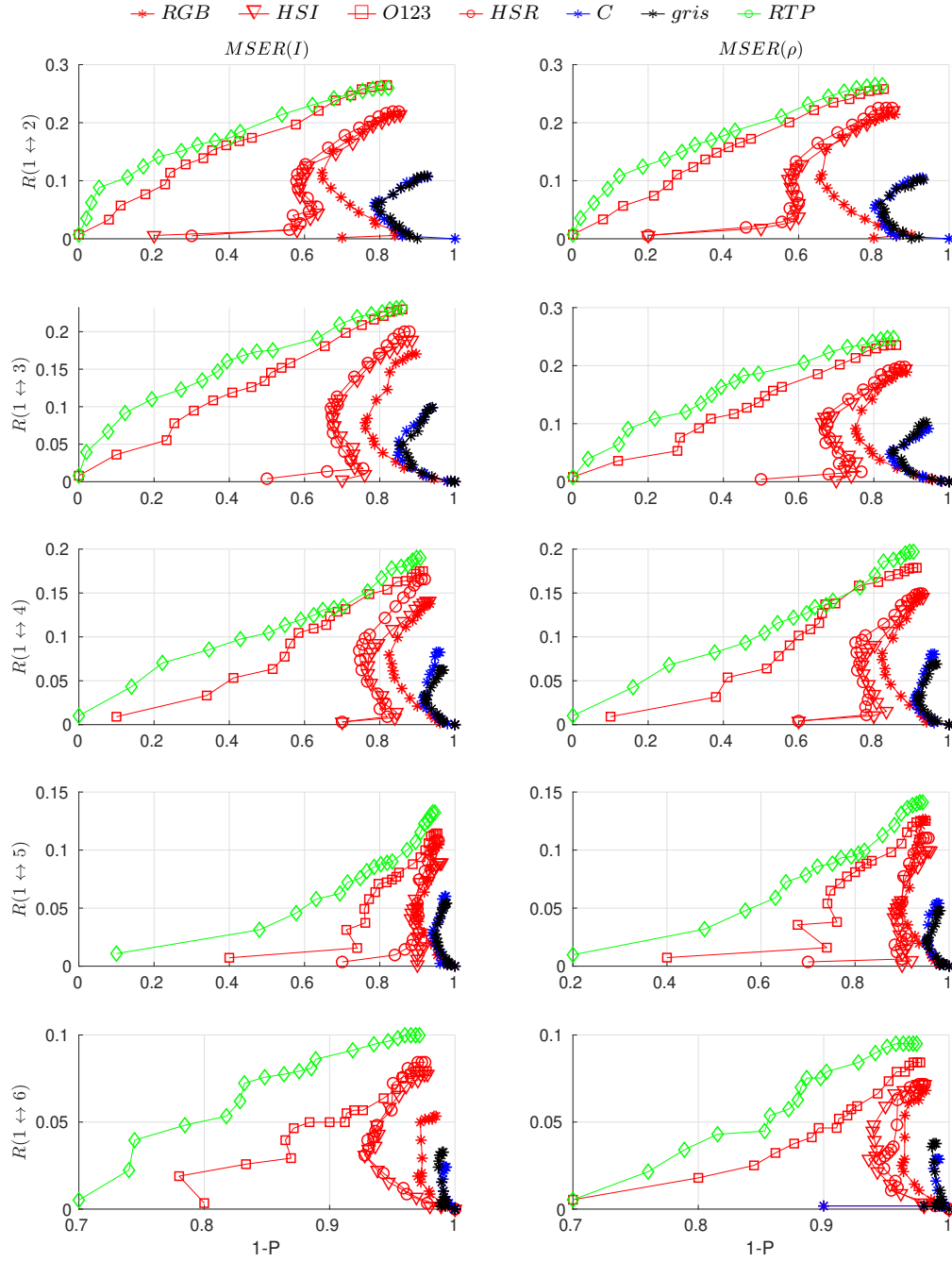


Figura B.2: R vs 1-P para el descriptor SURF en el conjunto Wall.

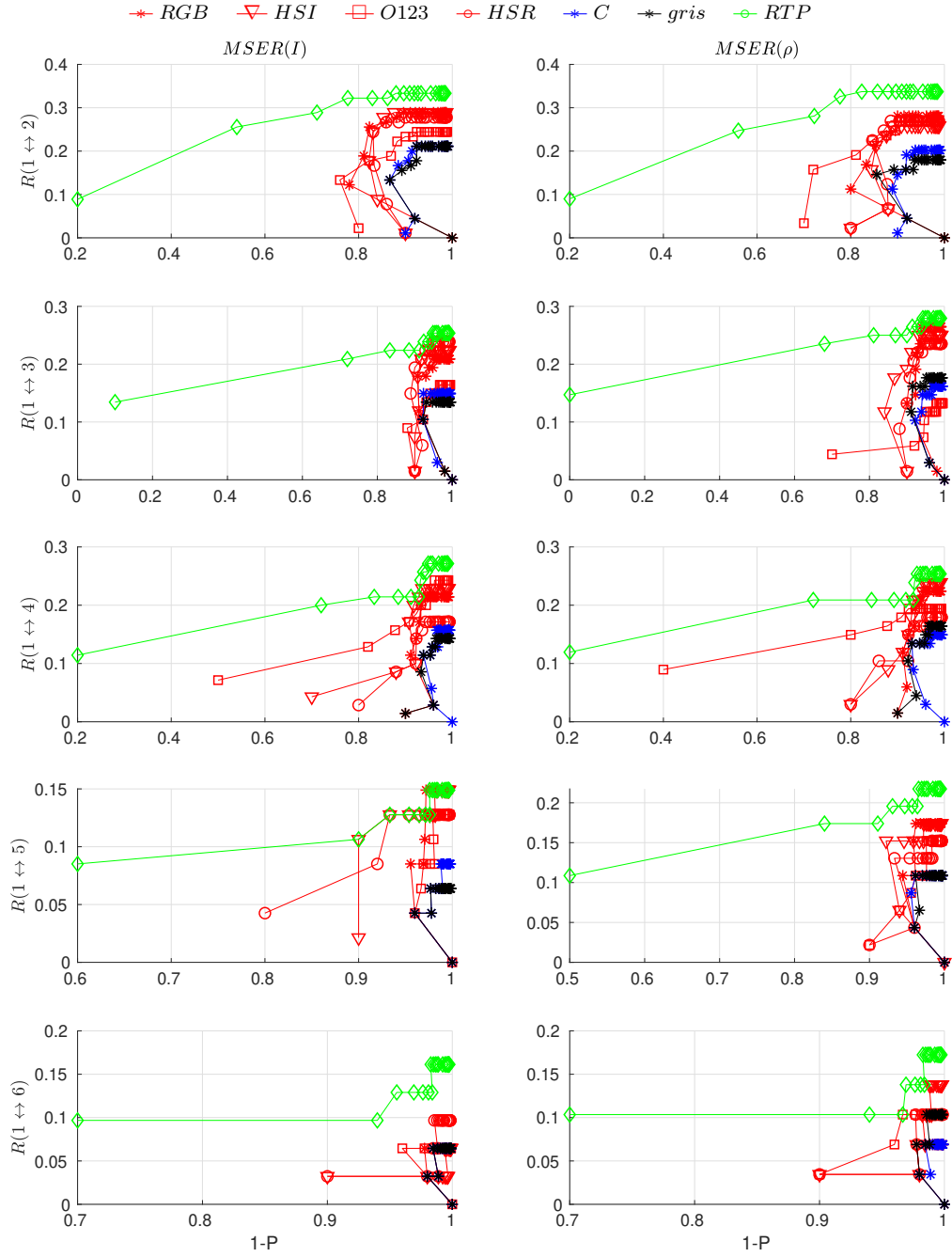


Figura B.3: R vs $1-P$ para el descriptor SURF en el conjunto Bark.

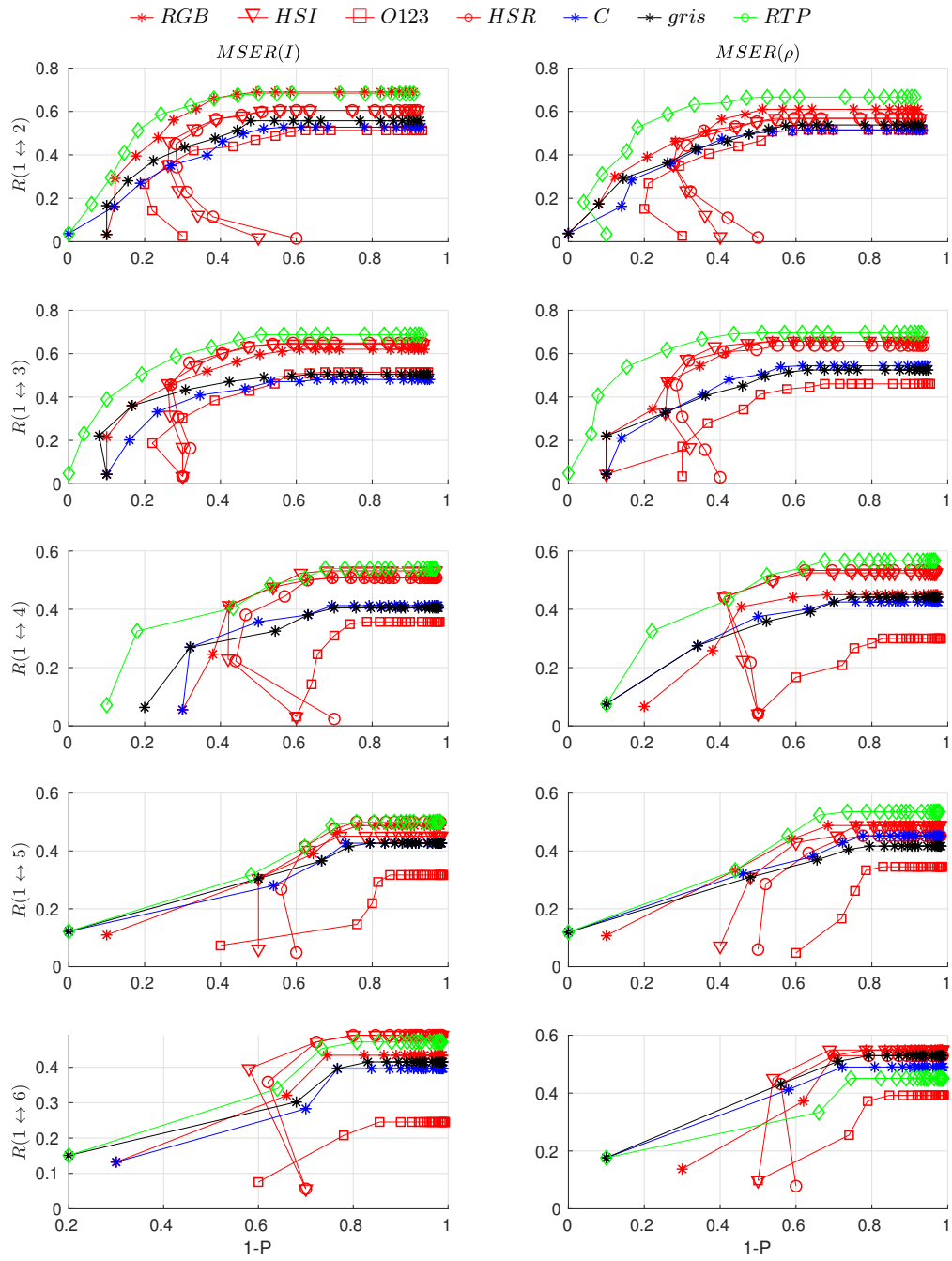


Figura B.4: R vs $1-P$ para el descriptor SURF en el conjunto Bikes.

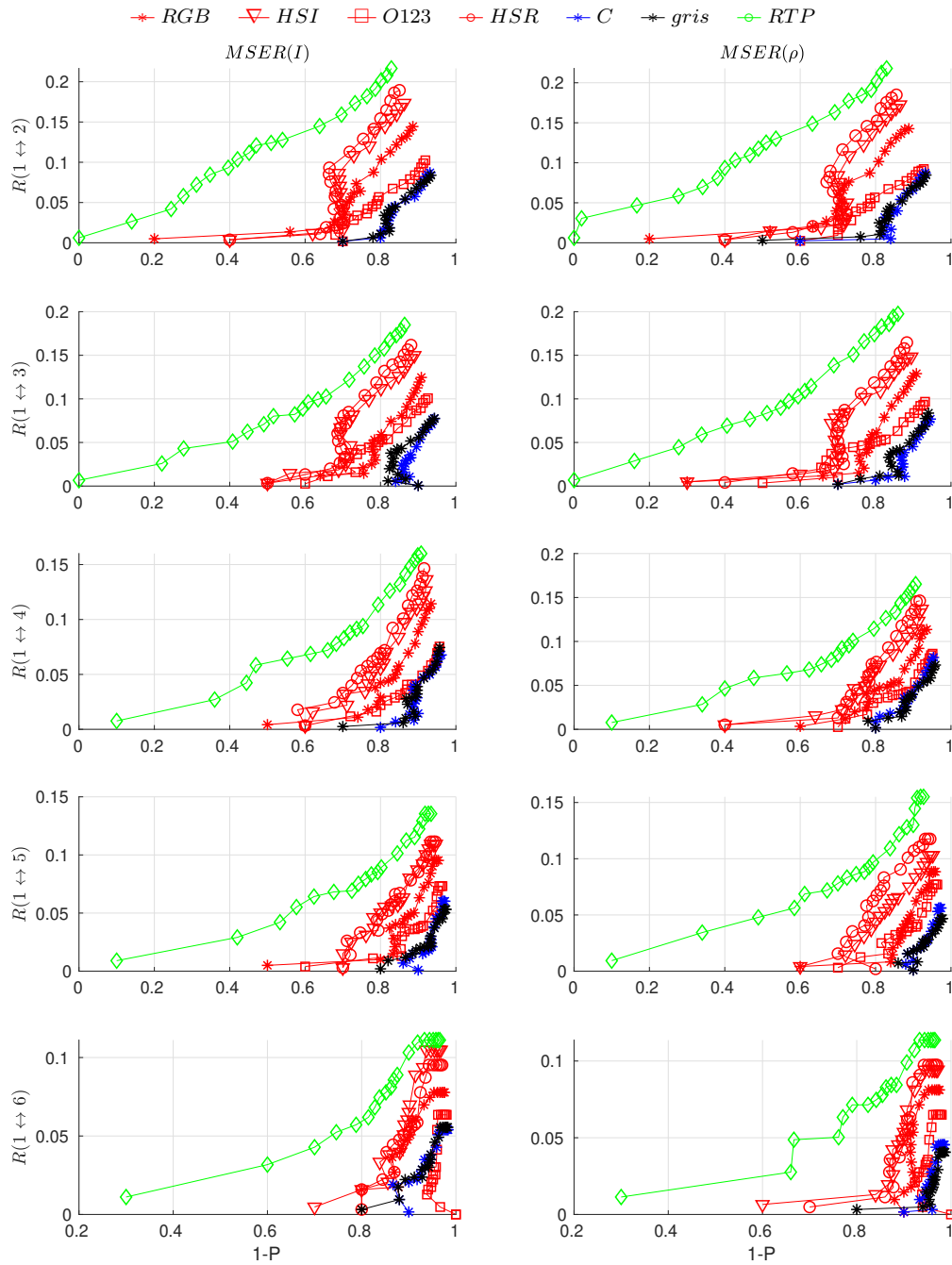


Figura B.5: R vs $1-P$ para el descriptor SURF en el conjunto Trees.

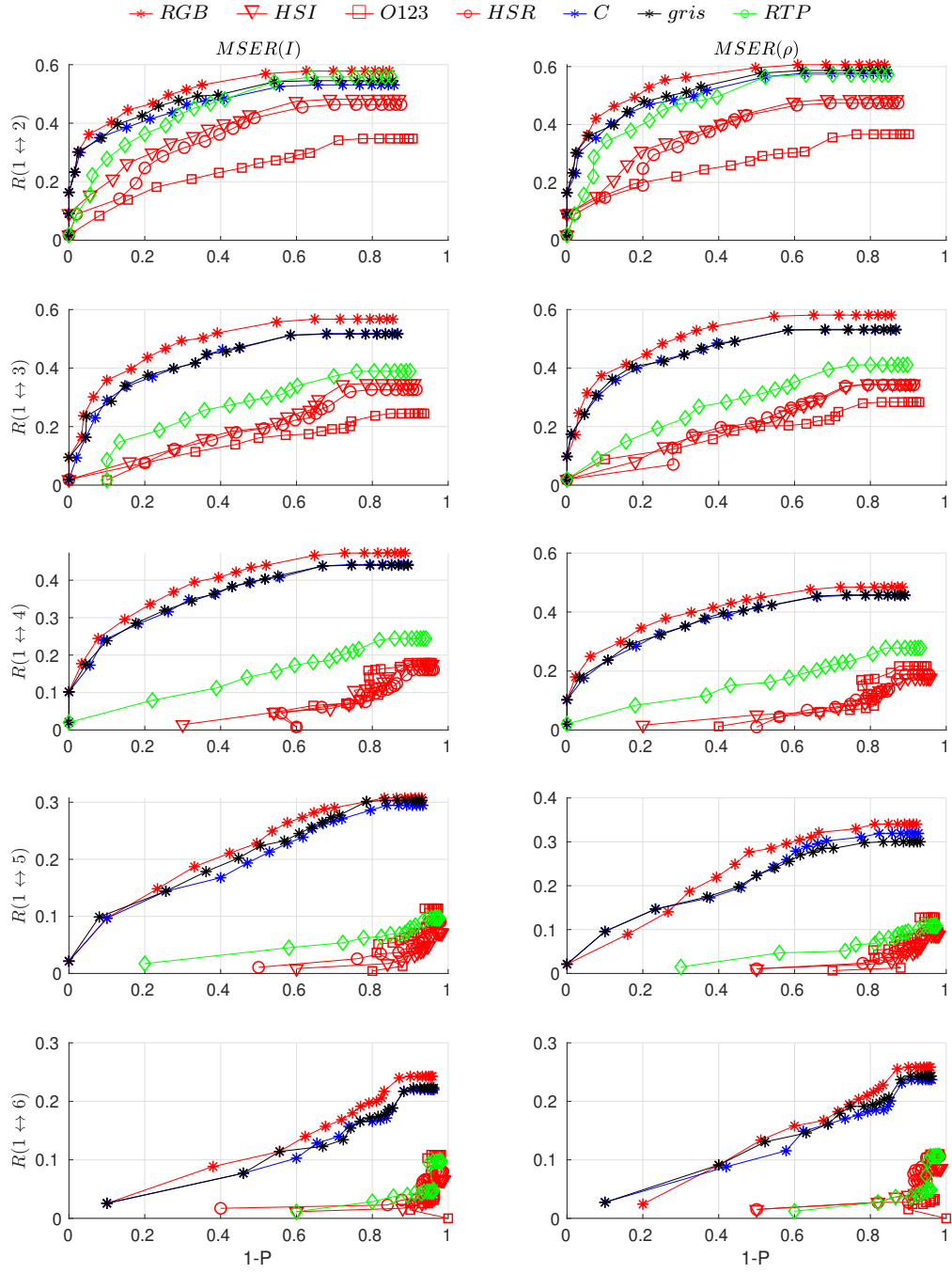


Figura B.6: R vs 1-P para el descriptor SURF en el conjunto UBC.

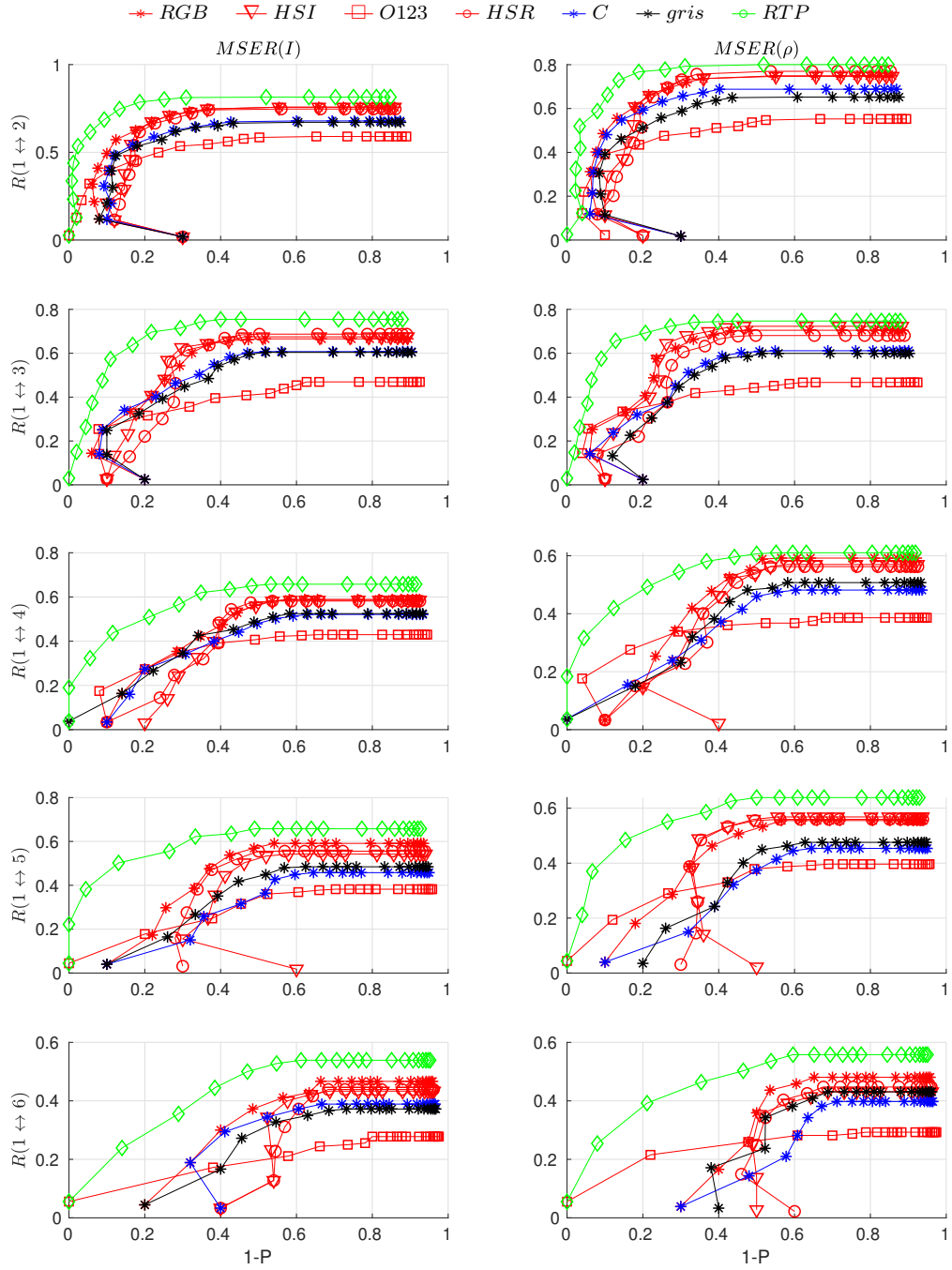


Figura B.7: R vs 1-P para el descriptor SURF en el conjunto Leuven.

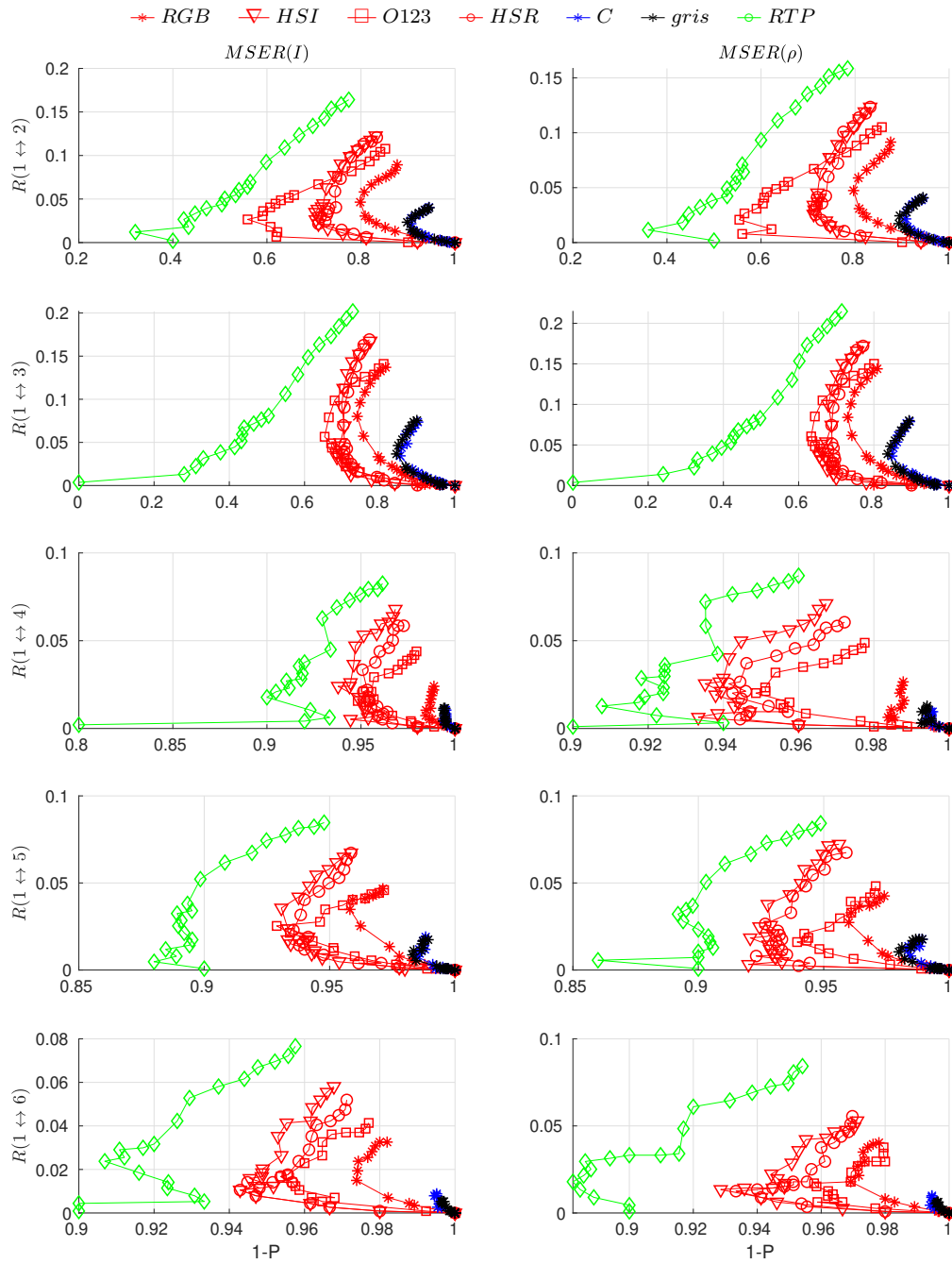


Figura B.8: R vs 1-P para el descriptor SURF en el conjunto Univalle.

Apéndice C

Curvas $F_{0.5}$ promedio en la etapa de descripción basada en SURF

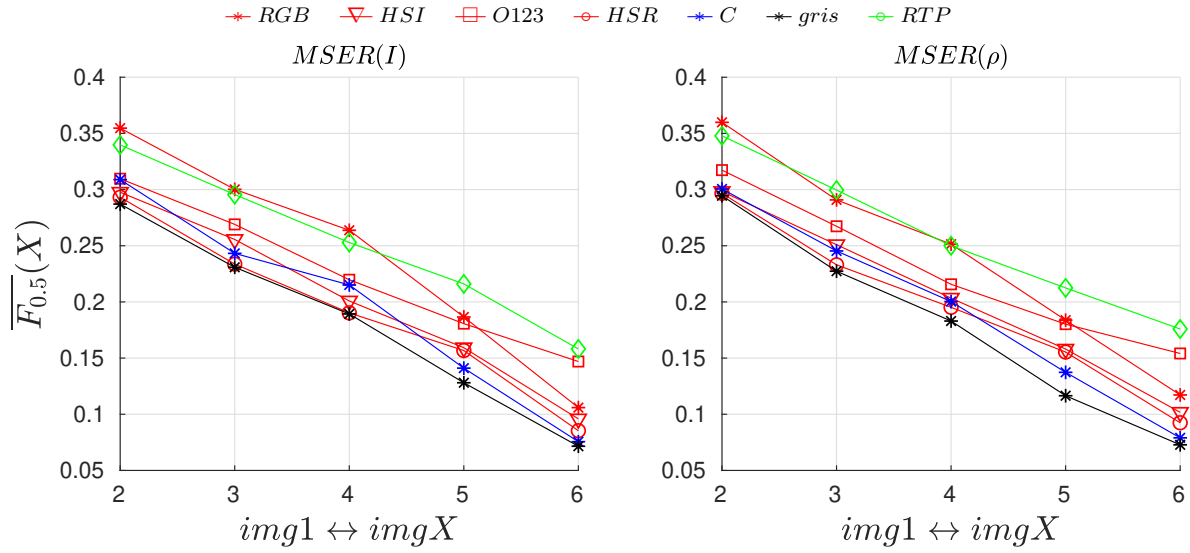


Figura C.1: F para cada par del conjunto Graffiti con el descriptor SURF con color

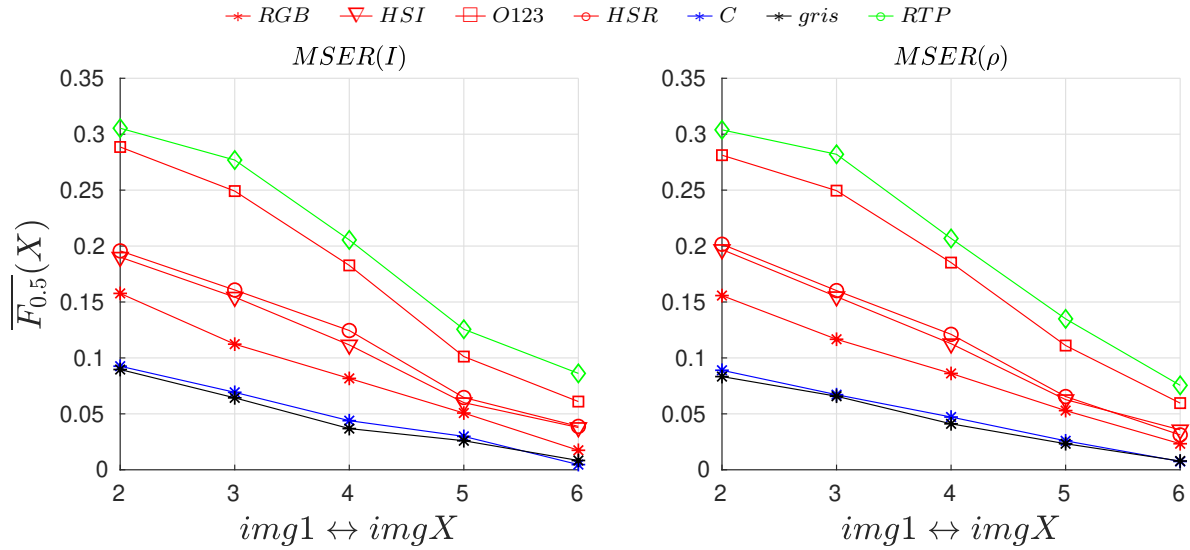


Figura C.2: F para cada par del conjunto Wall con el descriptor SURF con color

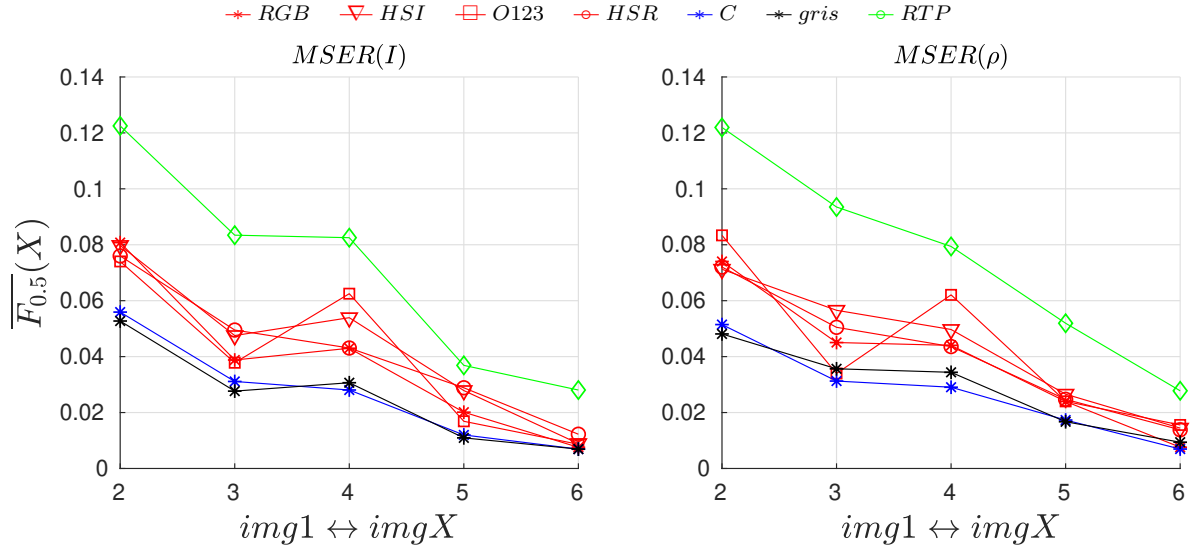


Figura C.3: F para cada par del conjunto Bark con el descriptor SURF con color

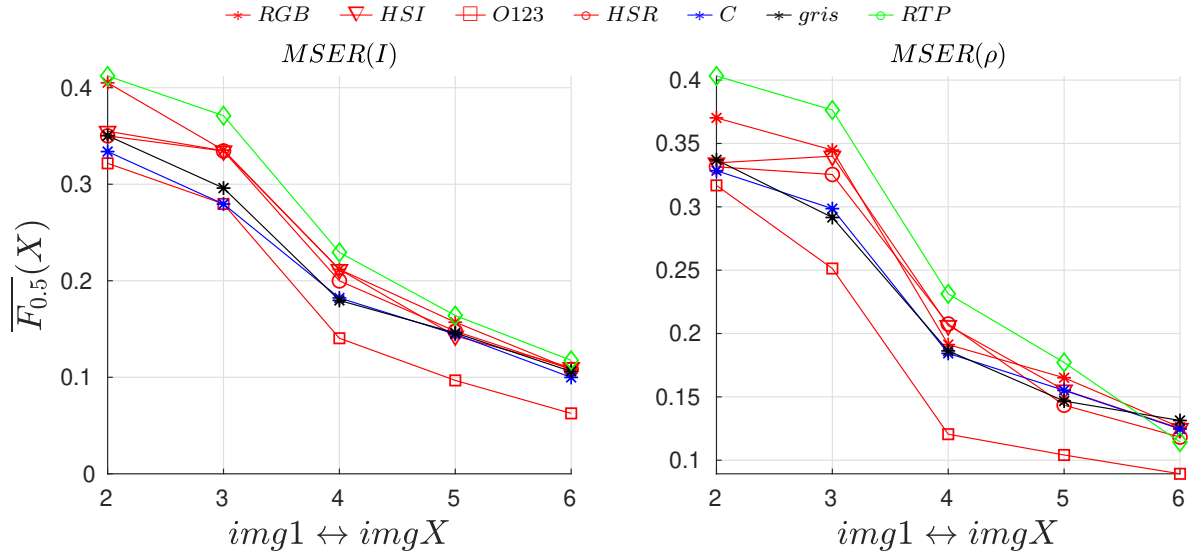


Figura C.4: F para cada par del conjunto Bikes con el descriptor SURF con color

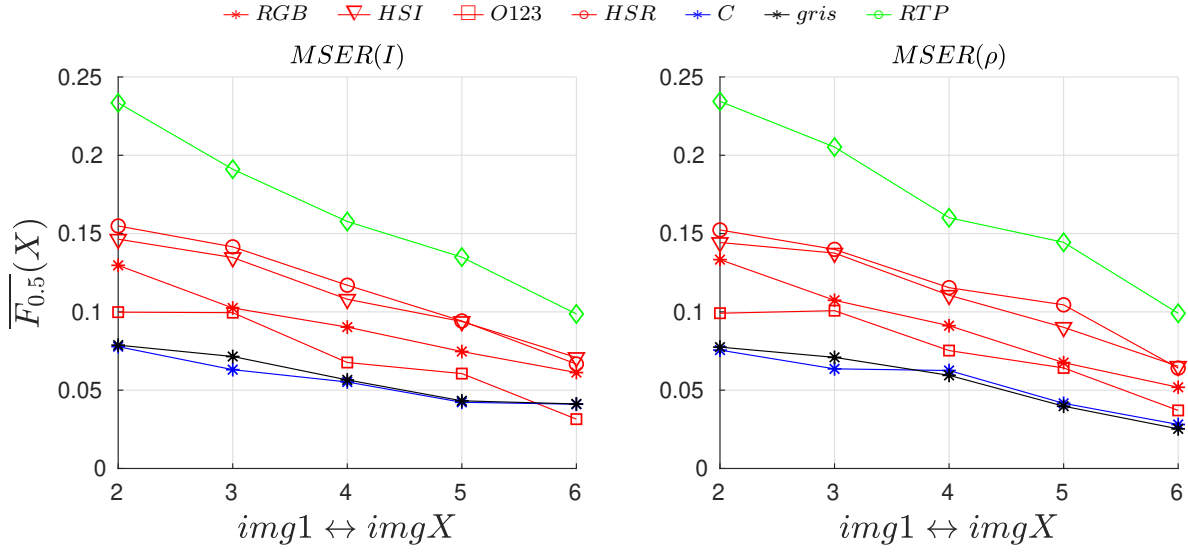


Figura C.5: F para cada par del conjunto Trees con el descriptor SURF con color

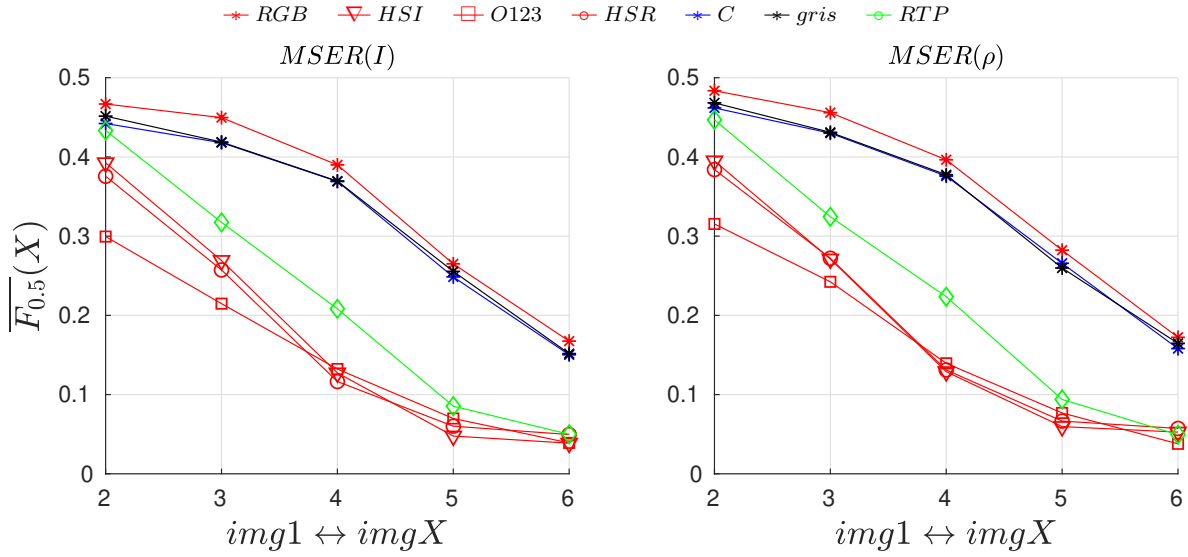


Figura C.6: F para cada par del conjunto UBC con el descriptor SURF con color

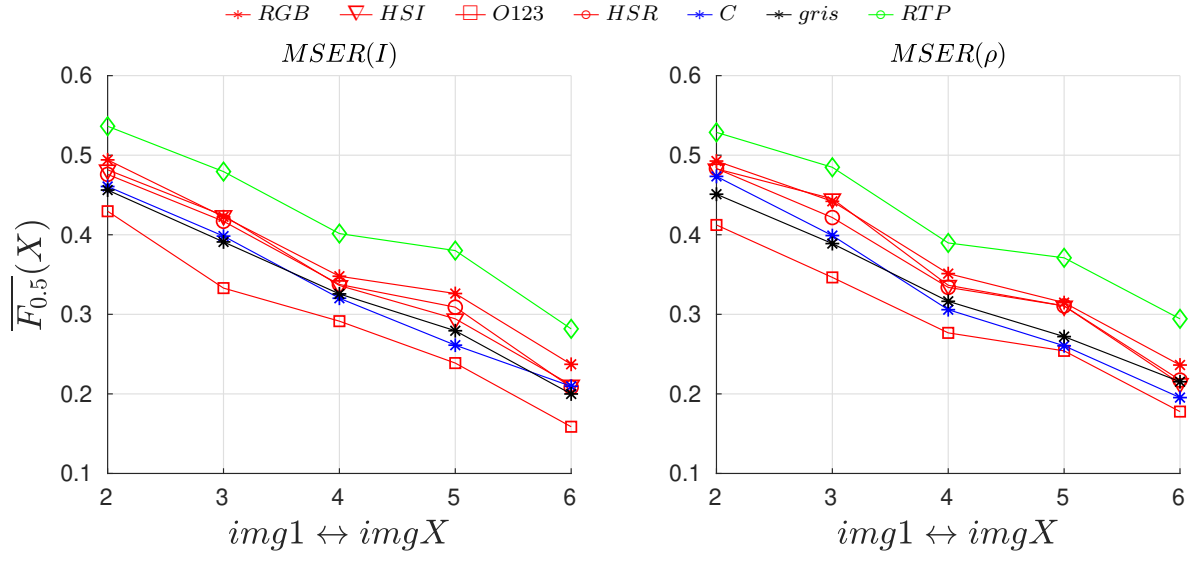


Figura C.7: F para cada par del conjunto Leuven con el descriptor SURF con color

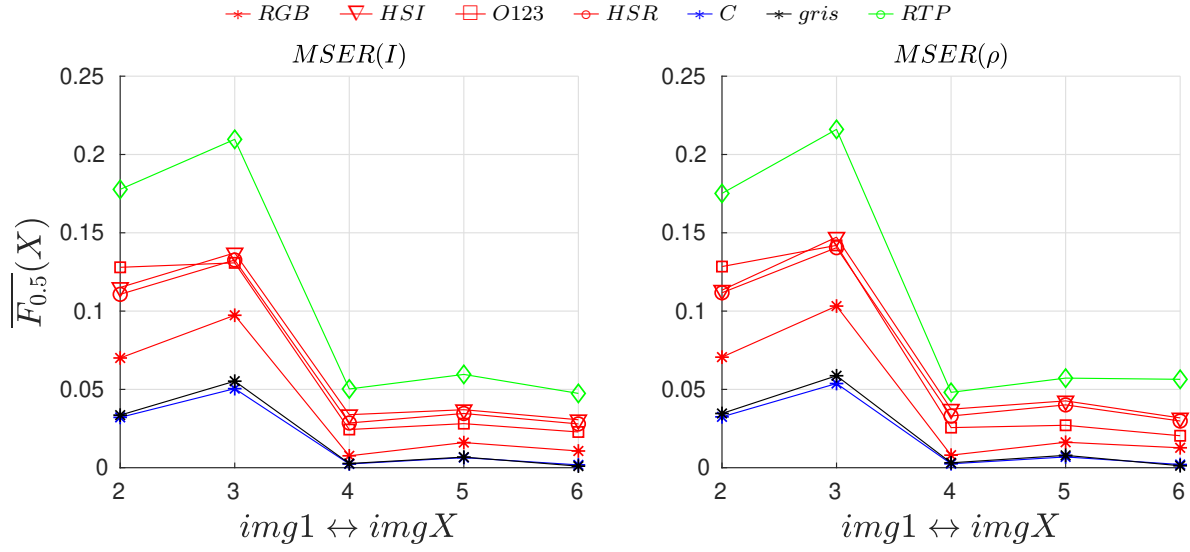


Figura C.8: F para cada par del conjunto Univalle con el descriptor SURF con color

Apéndice D

Curvas de Razón de aciertos vs 1-Precisión en la etapa de descripción basada en SIFT

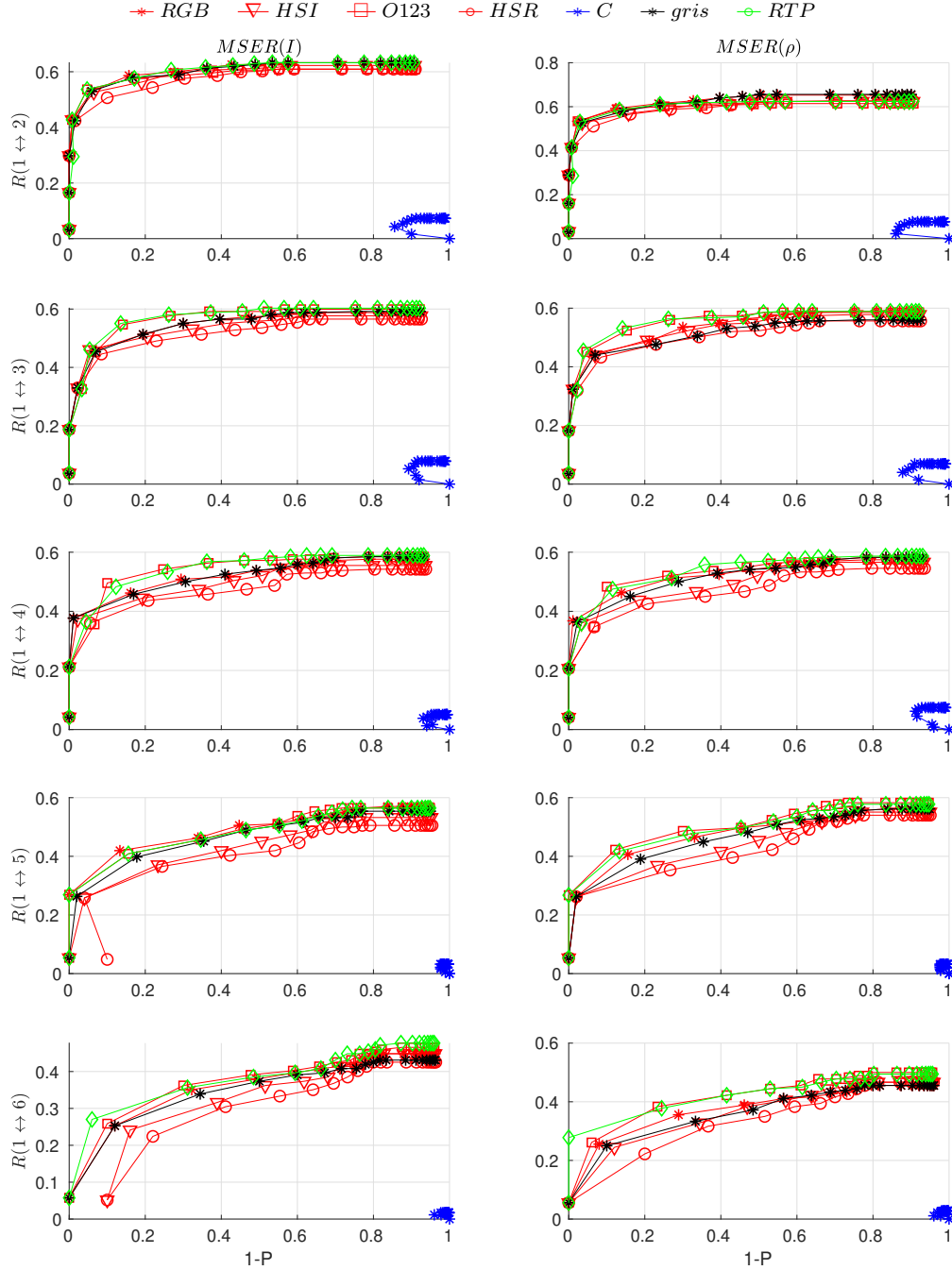


Figura D.1: R vs 1-P para el descriptor SIFT en el conjunto Graffiti.

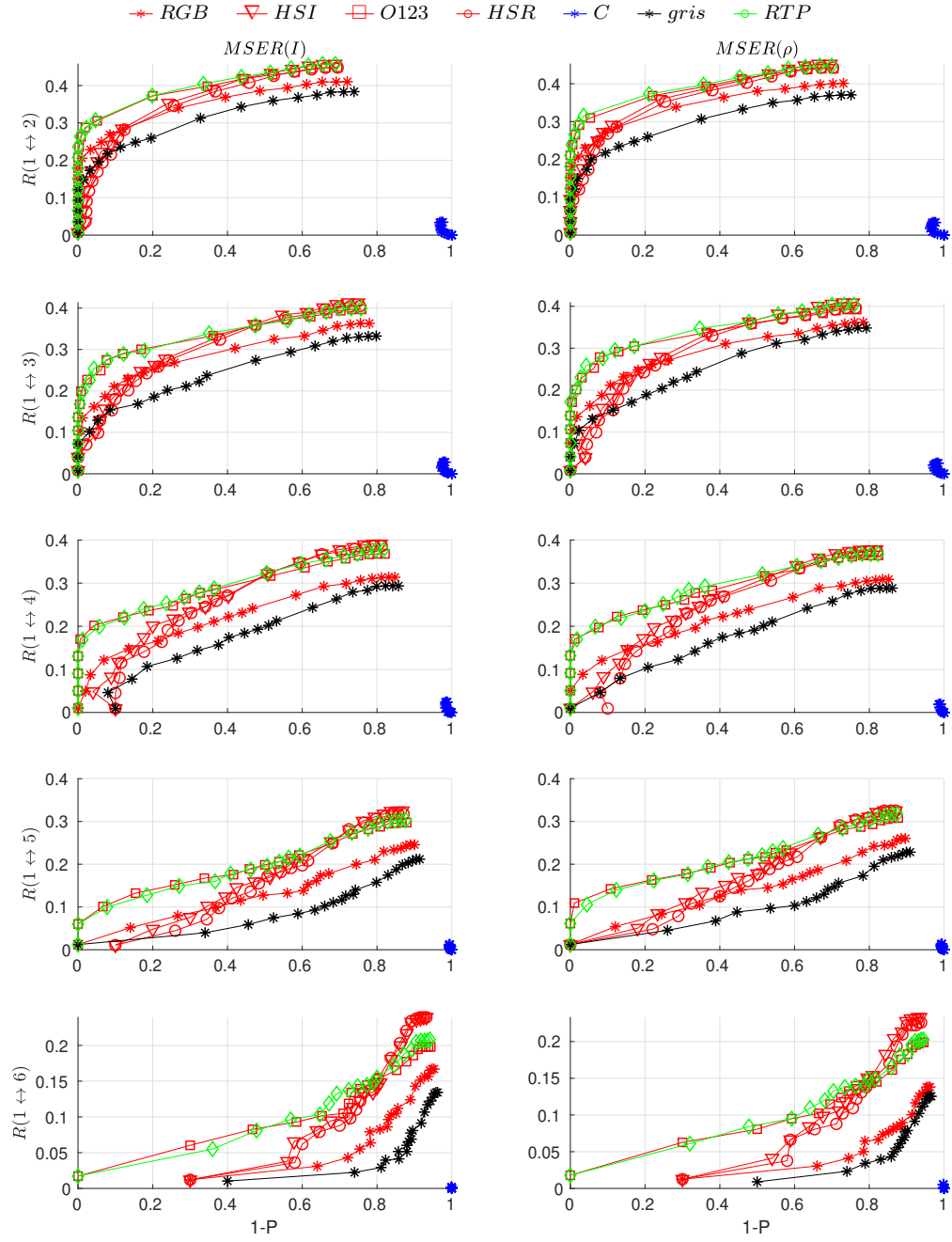


Figura D.2: R vs $1-P$ para el descriptor SIFT en el conjunto Wall.

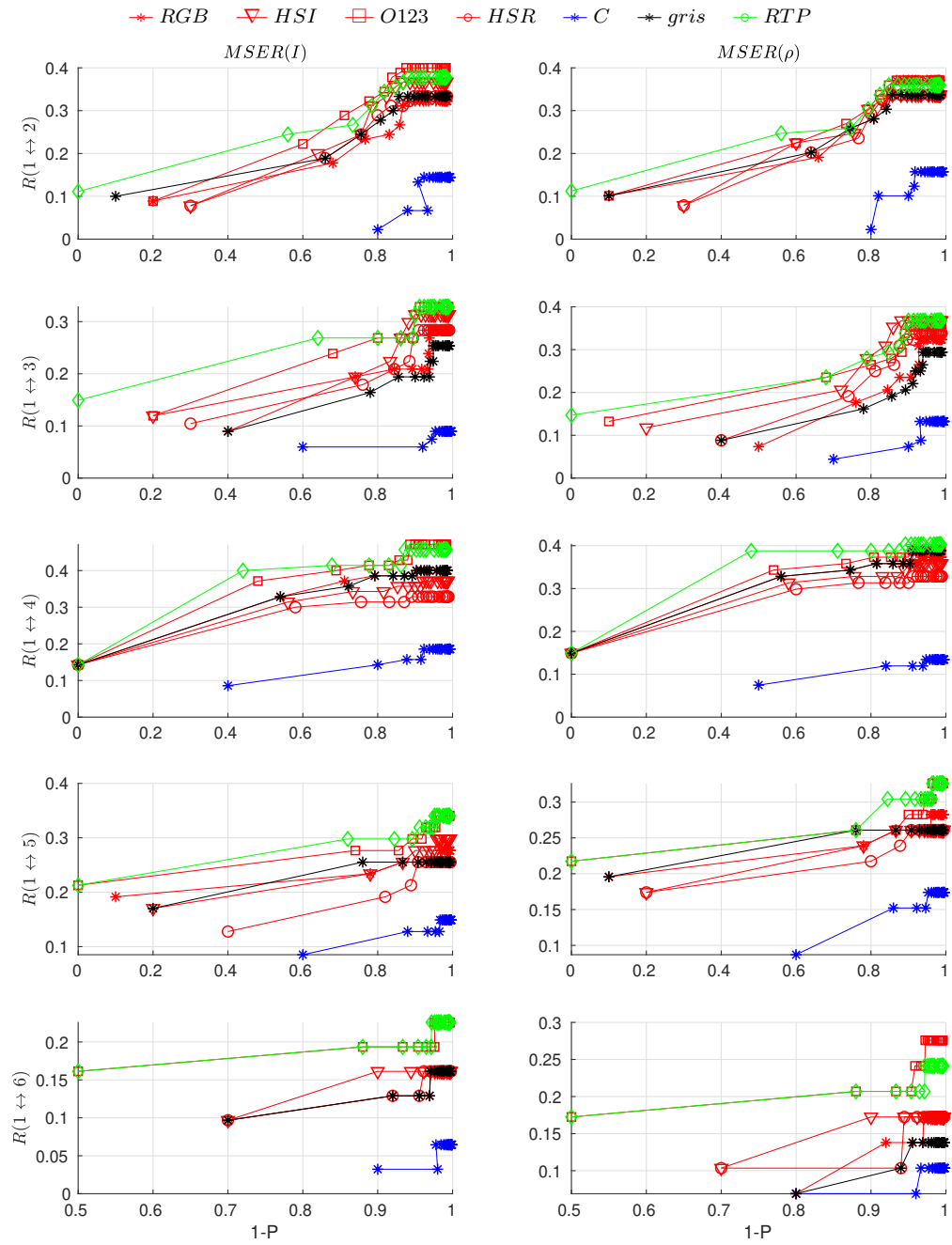


Figura D.3: R vs $1-P$ para el descriptor SIFT en el conjunto Bark.

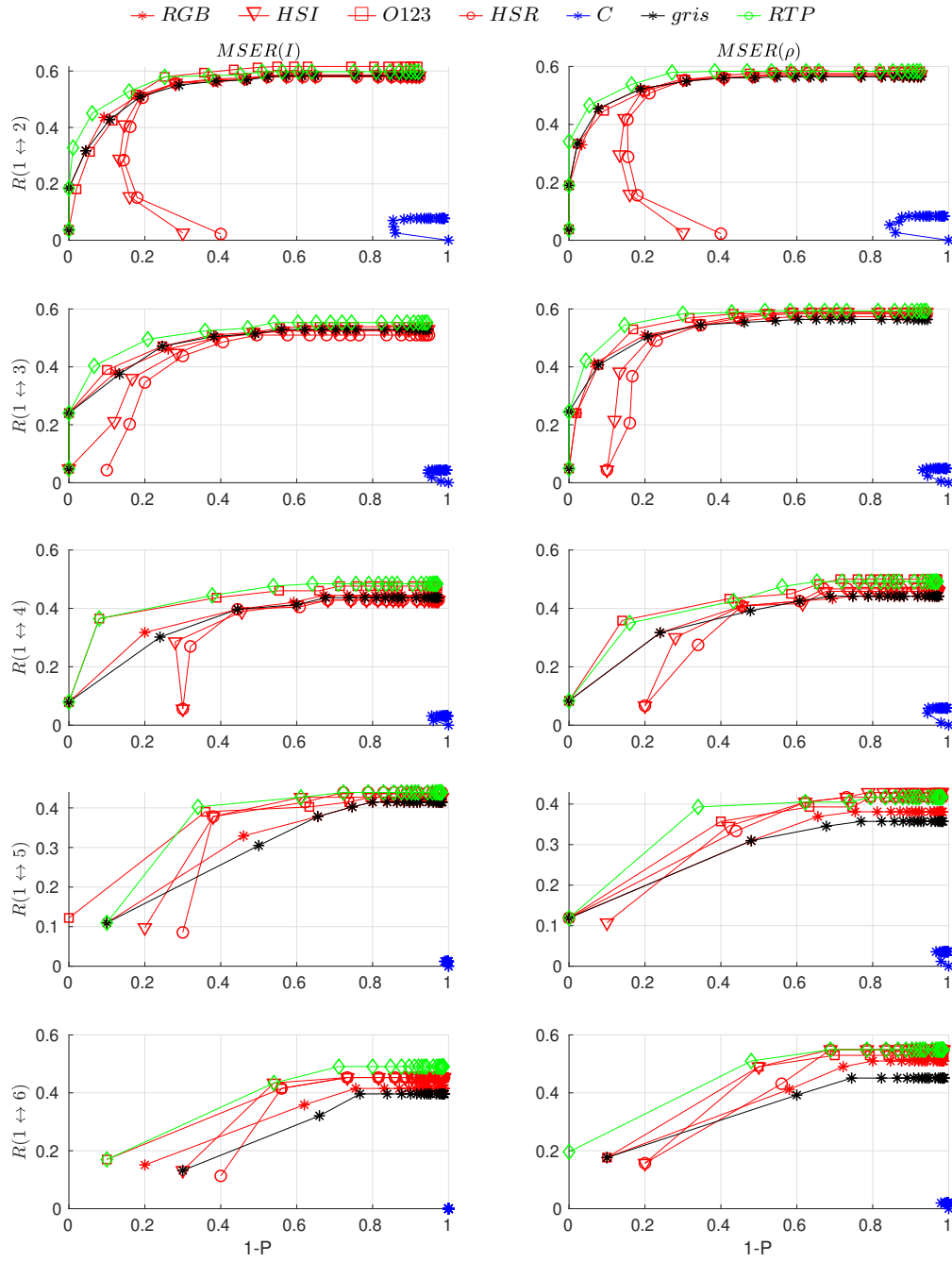


Figura D.4: R vs 1-P para el descriptor SIFT en el conjunto Bikes.

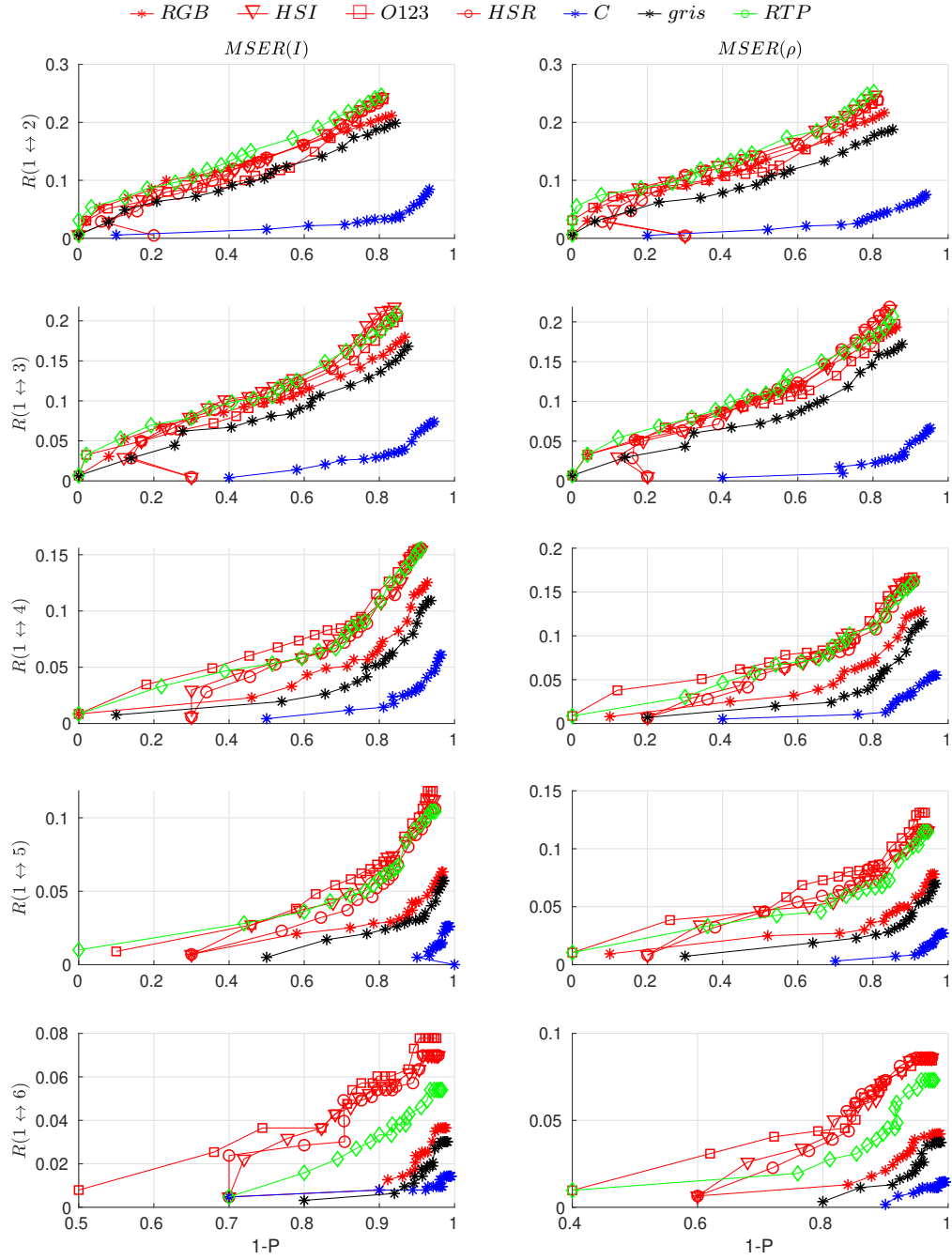


Figura D.5: R vs 1-P para el descriptor SIFT en el conjunto Trees.

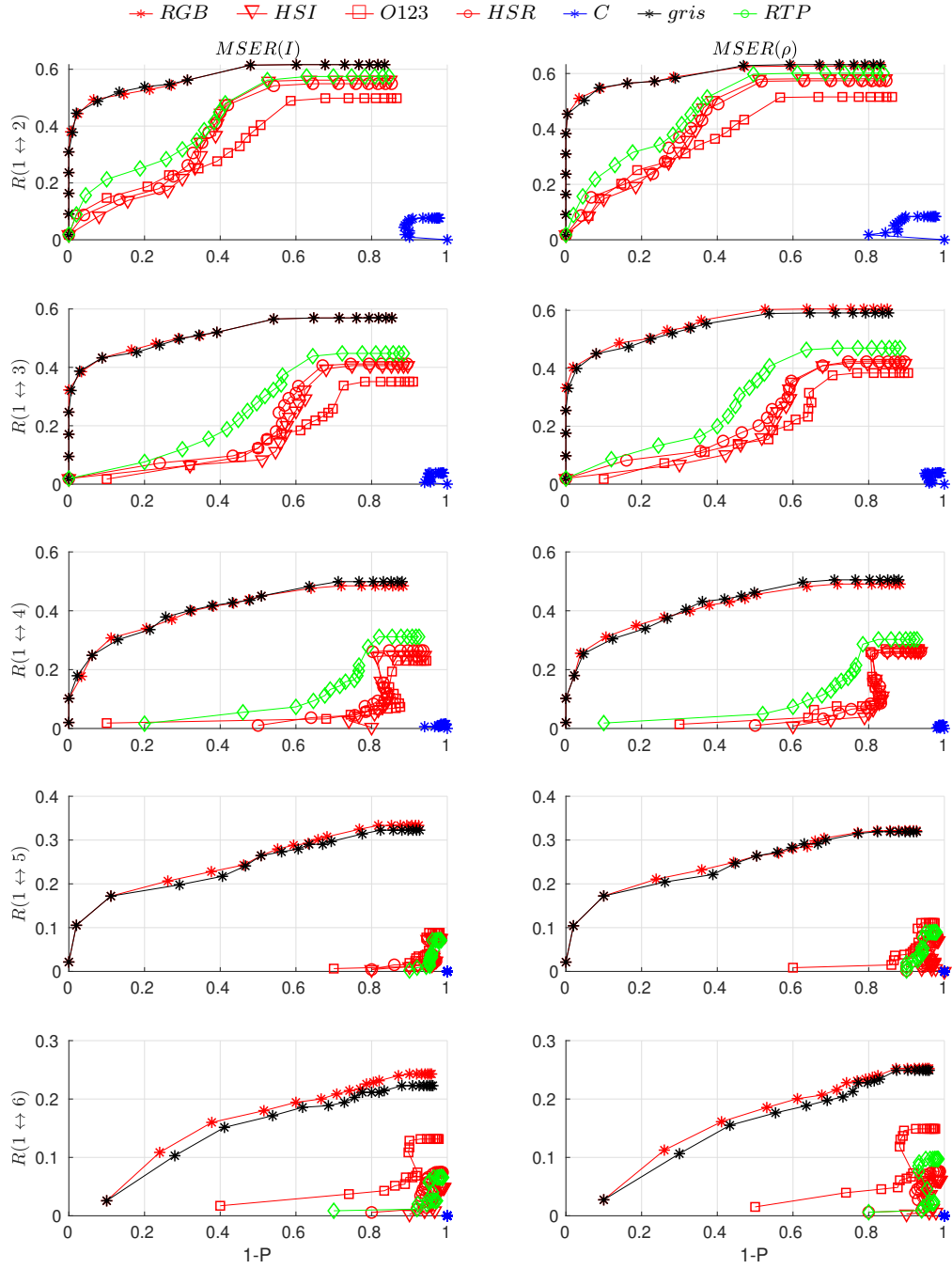


Figura D.6: R vs 1-P para el descriptor SIFT en el conjunto UBC.

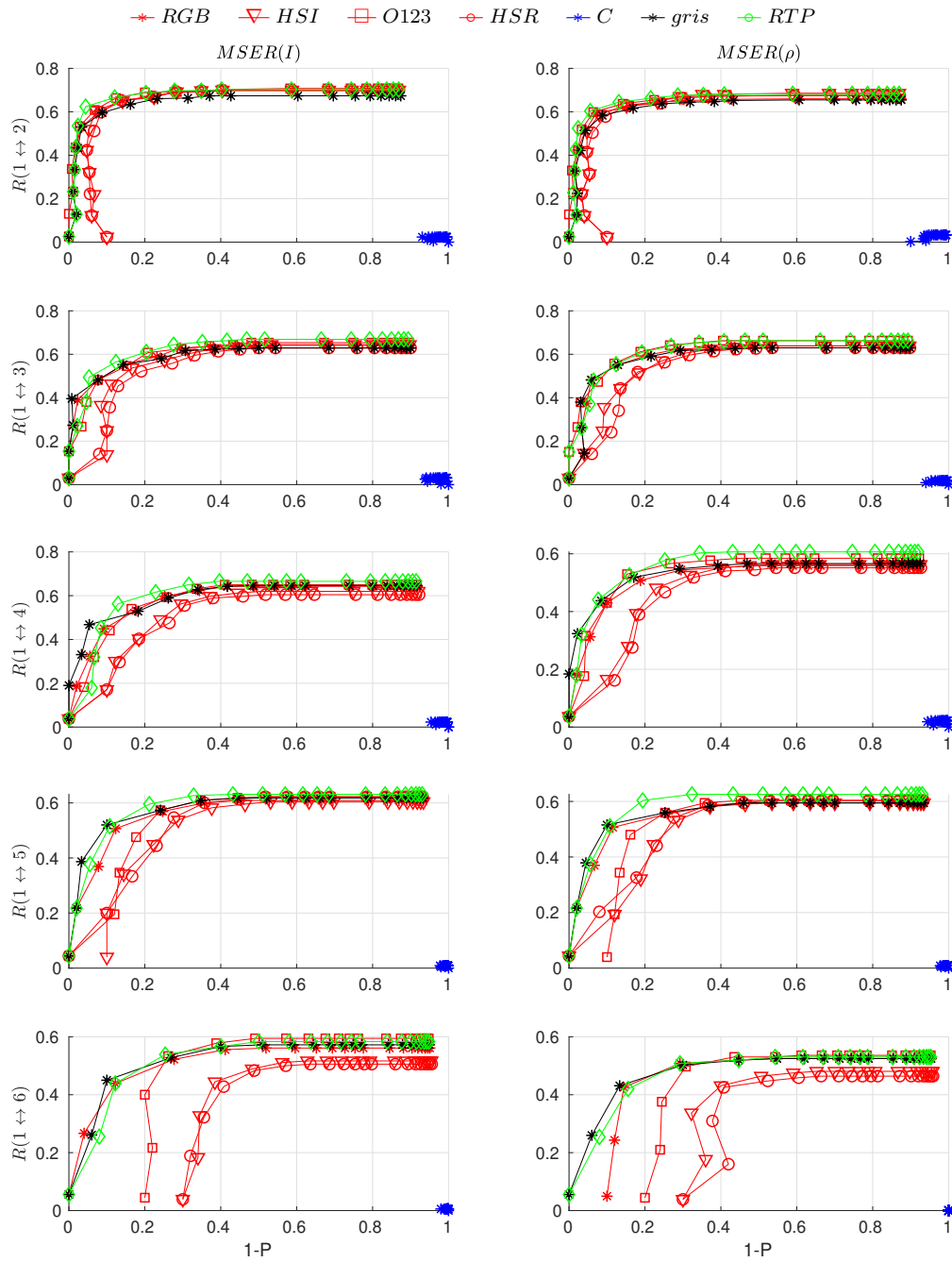


Figura D.7: R vs $1-P$ para el descriptor SIFT en el conjunto Leuven.

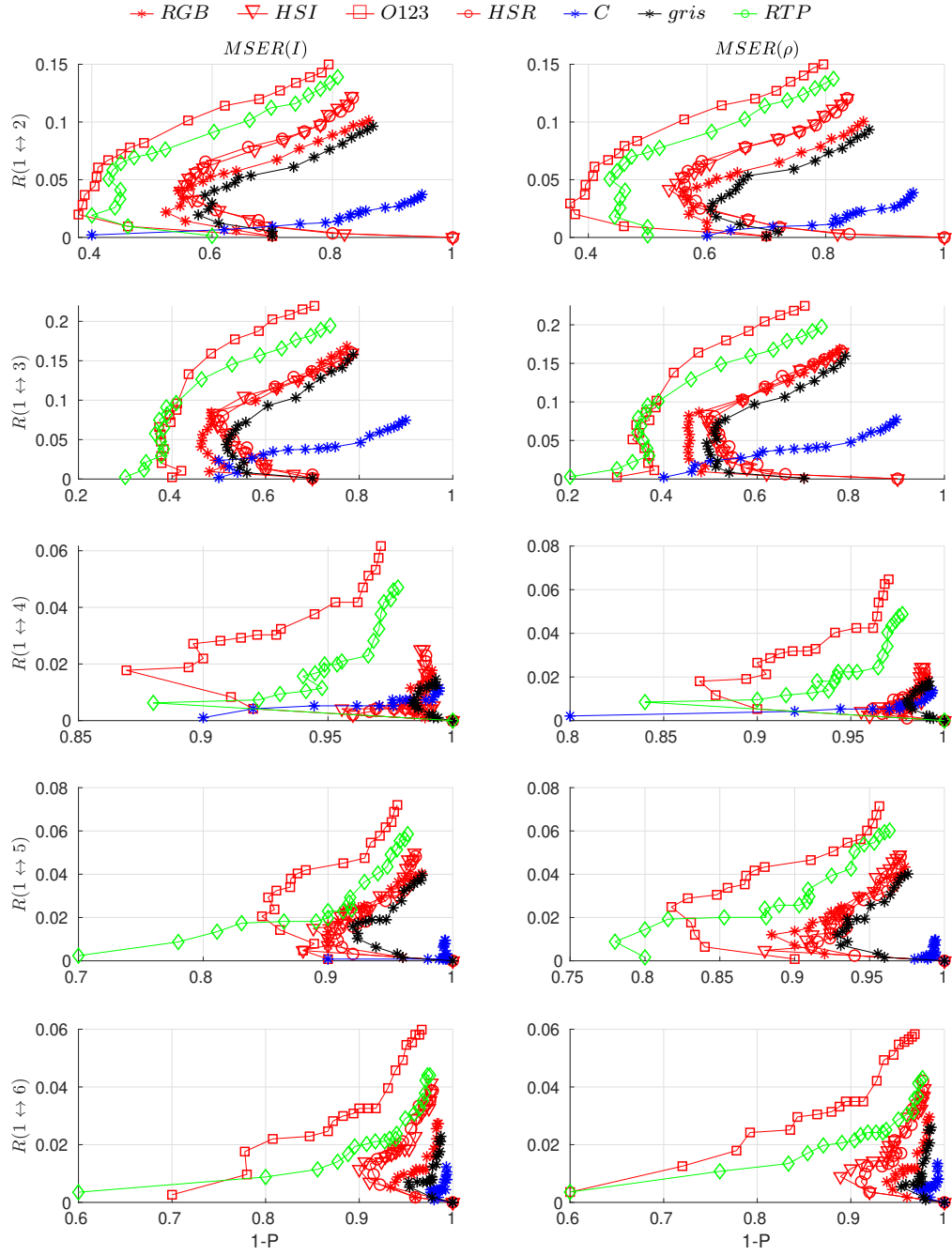


Figura D.8: R vs $1-P$ para el descriptor SIFT en el conjunto Univalle.

Apéndice E

Curvas $F_{0.5}$ promedio en la etapa de descripción basada en SIFT

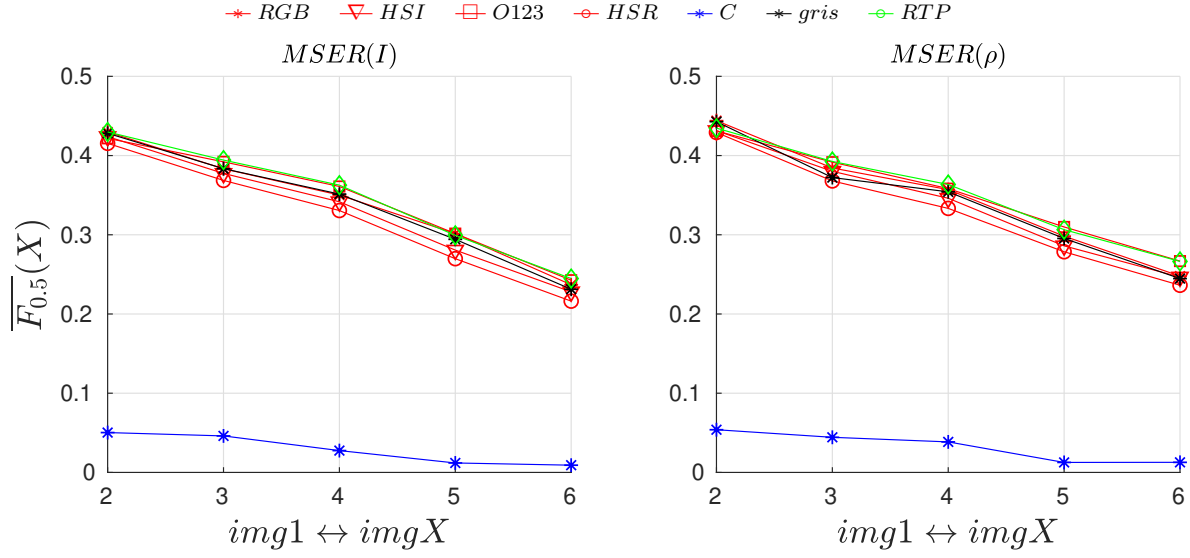


Figura E.1: F para cada par del conjunto Graffiti con el descriptor SIFT con color

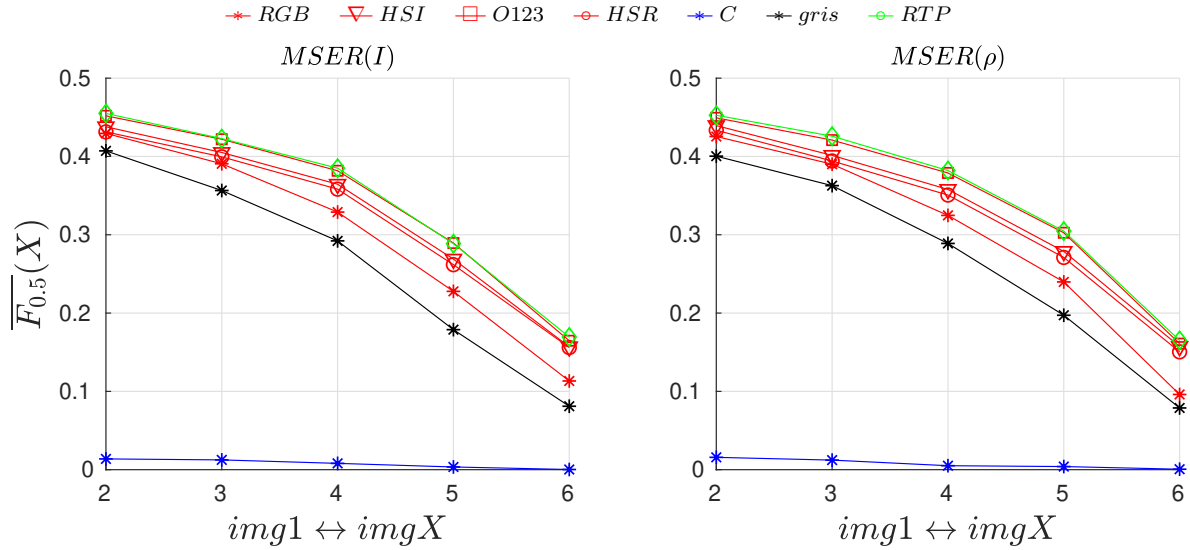


Figura E.2: F para cada par del conjunto Wall con el descriptor SIFT con color

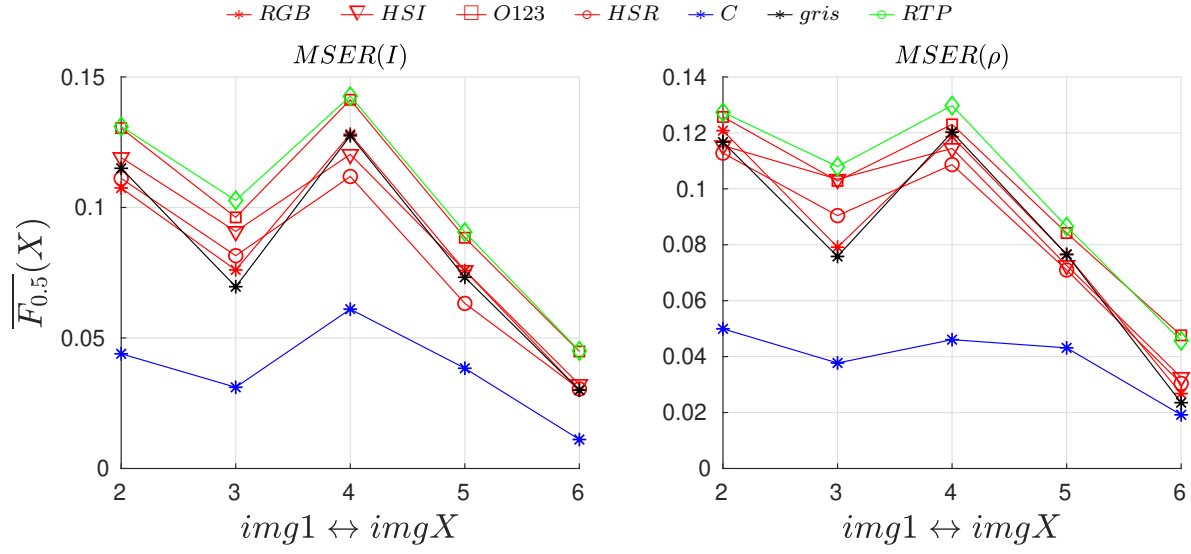


Figura E.3: F para cada par del conjunto Bark con el descriptor SIFT con color

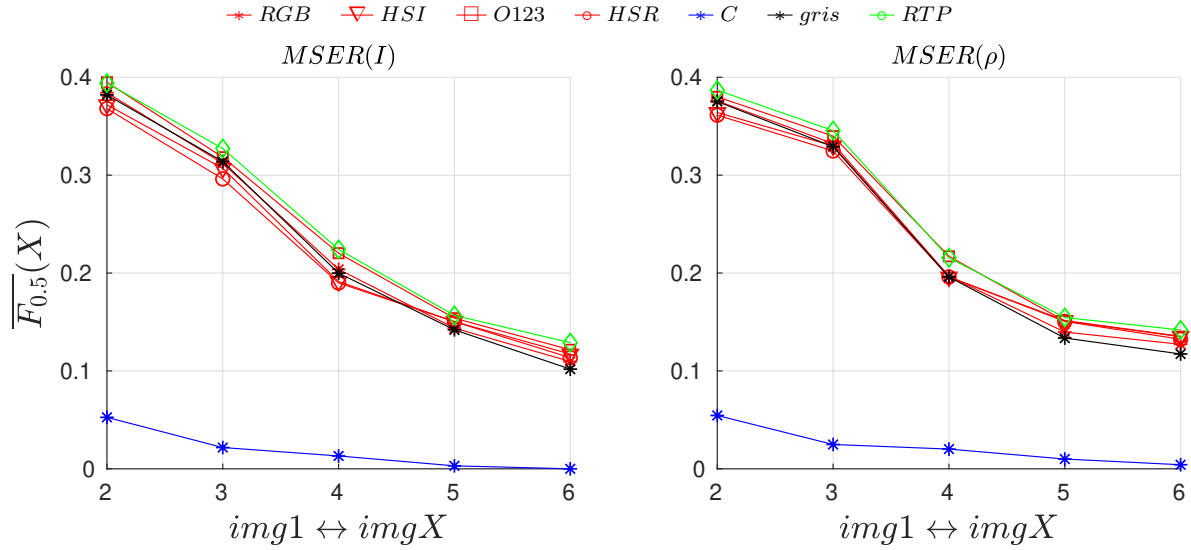


Figura E.4: F para cada par del conjunto Bikes con el descriptor SIFT con color

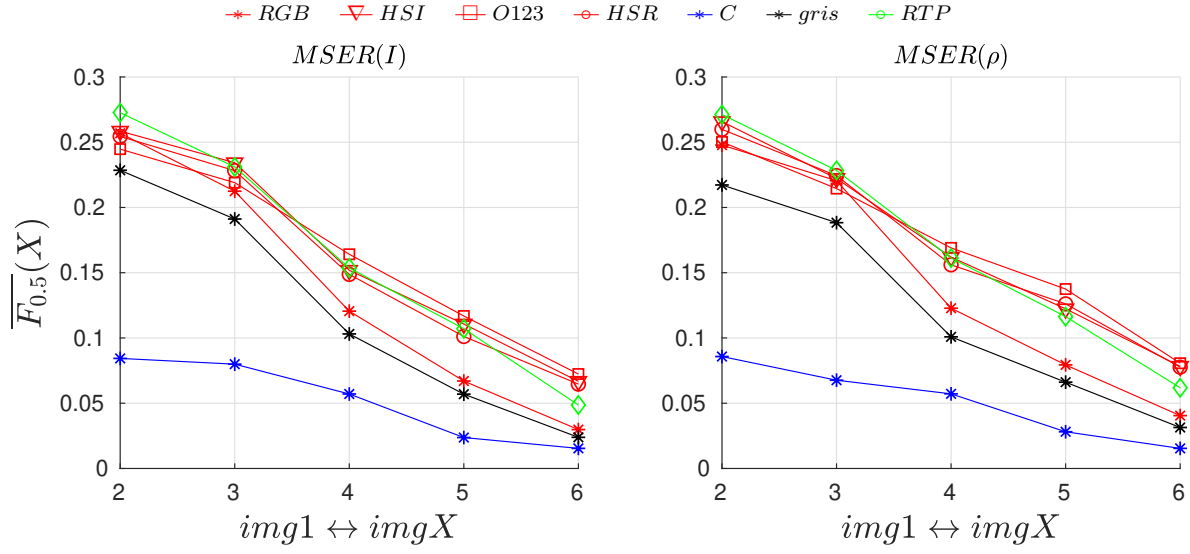


Figura E.5: F para cada par del conjunto Trees con el descriptor SIFT con color

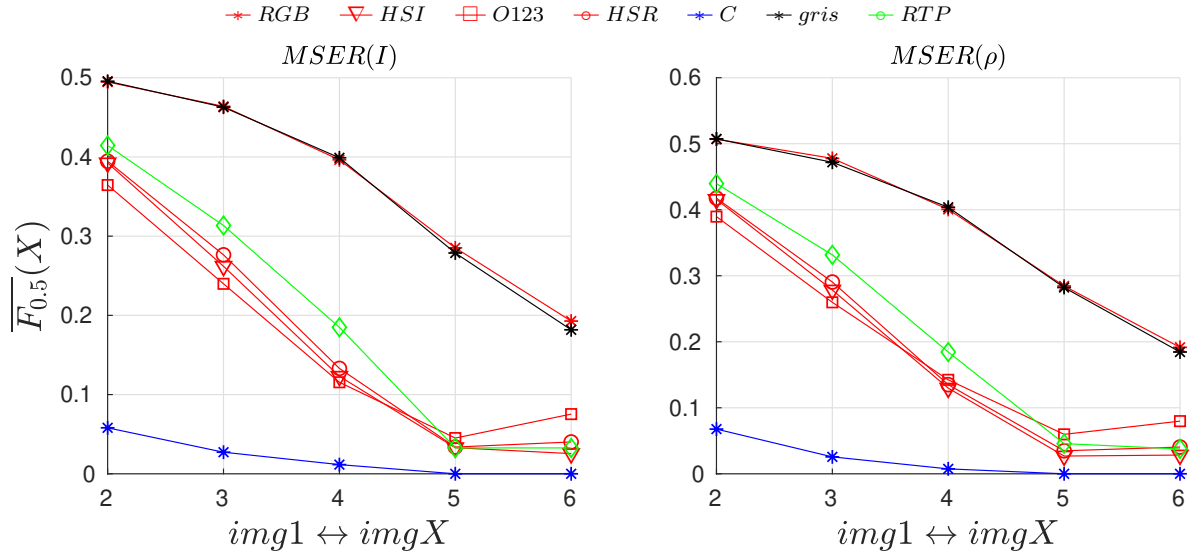


Figura E.6: F para cada par del conjunto UBC con el descriptor SIFT con color

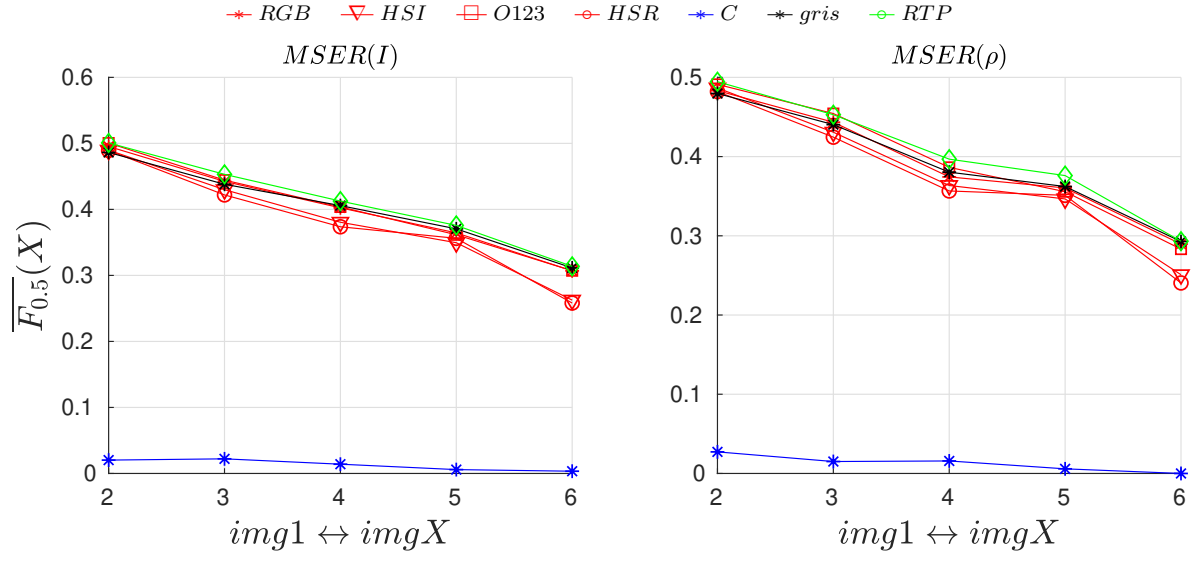


Figura E.7: F para cada par del conjunto Leuven con el descriptor SIFT con color

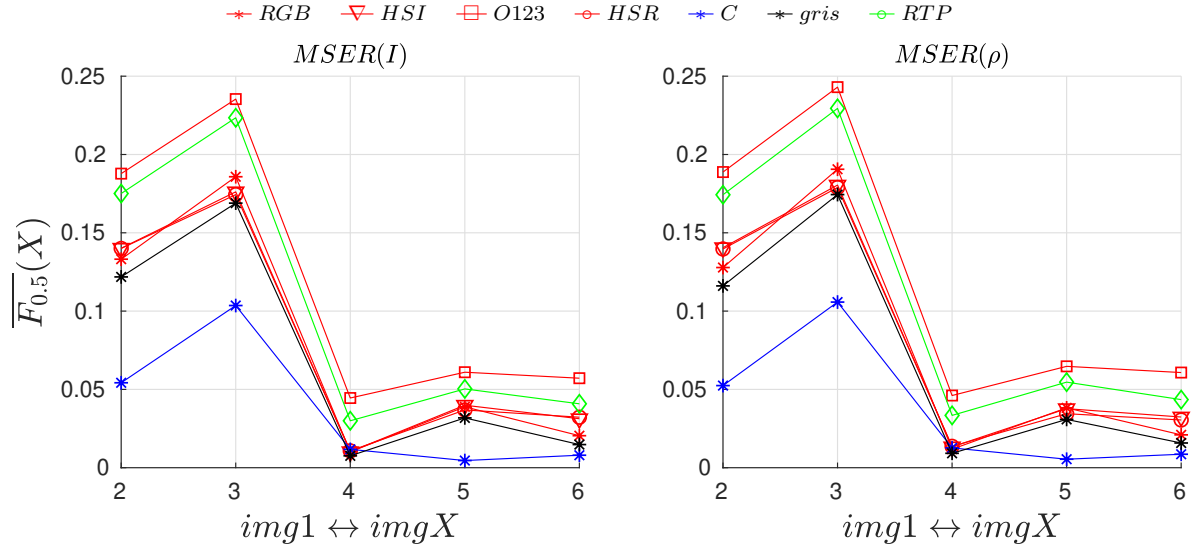


Figura E.8: F para cada par del conjunto Univalle con el descriptor SIFT con color